SUPPLEMENTO
AL VOLUME XVII, SERIE X, DEL
NUOVO CIMENTO
A CURA DELLA SOCIETÀ ITALIANA DI FISICA

1960 3° Trimestre N. 1

Quantum Mechanics and the Uniqueness of the World.

B. BERTOTTI (*)
Project Matterhorn, Princeton University - Princeton, N. J.

(ricevuto l'11 Gennaio 1960)

1. – Quantum mechanics replaces the unique evolution of a classical physical system, fixed deterministically by the initial conditions, with a family of evolutions, to which a complex measure is attached. This infinite multiplicity of descriptions, which adequately meets the unpredictability of microscopic phenomena, on the other hand comes up against the basic datum of everyday life, *there is only one world*. From whatever point of view they are looked upon, the characteristic feature of the data upon which science is based is that they are there, unavoidably unique; and we cannot say that we understand them until we are able to characterize them against all the other possibilities.

The experimental physicist—and the cultivated layman—usually does not worry much about this. All that quantum mechanics can tell us, is often explained, is only a definite probability distribution for any physical quantity; the experiment itself will show which one of the infinite number of possible alternatives is « true ». He is told that this awkward situation arises from the inescapable and unpredictable perturbation produced by the measuring apparatus on a microscopic system, thereby tacitly or unknowingly implying that the present theory is only an overall and coarse description of what is going on and does not pay any attention to the actual, detailed mechanism of such a perturbation. As VON NEUMANN and others often pointed out [1], such a view

(*) This paper was partially written while the author was at the Institute for Advanced Study, Princeton, N. J.

[1] J. VON NEUMANN: *Mathematical Foundations of Quantum Mechanics* (Princeton, N. J., 1955), especially Chap. VI.

is contrary to the spirit and the letter of quantum mechanics—a *complete* and *self-consistent* theory. The fact that the physical observables do not have sharp values has nothing to do with the smallness of the action involved in the measurement, but is a *necessary* feature of any system, be it big or microscopic, be it isolated or not. The process of singling out a unique value—the observed value—of a physical quantity in a laboratory experiment—thereby re-establishing «the unique world»—has, according to the present theory, nothing to do with the instrument and transcends the domain of our knowledge. The measuring device is in fact itself a (big) *quantum mechanical* system, obeying the ordinary Schrödinger equation and undergoing quantum dispersion. If a particle in the form of a plane wave impinges on two counters, the state of the latter goes over into the superposition of two different states, in each of which one counter has responded, the other one has not. The square of the absolute value of the amplitude of each state gives the frequency of each alternative when the experiment is repeated over and over again; but it is in principle impossible to predict which one will respond in a definite case. Quantum mechanical alternatives must still be present in any macroscopical process, as long as it is ruled by ordinary quantum laws. How are we to understand the fact that, just by *looking* at the apparatus, we are able to make the decision? It is easy for a layman to shake his head skeptically in the obvious answer: of course quantum mechanics is only a statistical theory, analogous to classical thermodynamics, and, in due time, will be replaced by a detailed description. The failure of many outstanding physicists who attempted to open a new path in this direction for thirty years and the widespread indifference among the scientists in regard to this problem show the naïvety of such hope and suggest great caution.

2. – For a better understanding of the quantum theory of measurements it is in order to recall the meaning of the work done by G. LUDWIG [2].

Consider a plane wave impinging on a screen with two holes. We know [3] that if a particle is detected right behind one of them, the interference pattern is destroyed and the two partial waves propagate incoherently. How are we to understand the destruction of phase interference which occurs in any physical measurement? Does this effect imply a sharp value for the observable?

In von Neumann's notation [4] we are concerned with the passage from a statistical matrix U to the matrix

$$U' = \sum P_\varepsilon U P_\varepsilon \,. \tag{1}$$

---

[2] G. LUDWIG: *Zeit. f. Phys.*, **135**, 483 (1953); **150**, 346 (1958); **152**, 98 (1958).

[3] A. ADAM, L. JÁNOSSY and P. VARGA: *Acta Phys. Hungarica*, **4**, 301 (1955); E. BRANNEN and H. I. S. FERGUSON: *Nature*, **178**, 481 (1956); R. Q. TWISS and A. G. LITTLE: *Aust. J. Phys.*, **12**, 77 (1959).

[4] G. LÜDERS: *Ann. Phys.*, **8**, 322 (1956).

where P_ε is the projection operator on the eigenmanifold V_ε of the quantity we have observed. The ordinary time evolution

$$U \to \exp[-iHt]\, U \exp[iHt] \tag{2}$$

must be sufficient to generate such a transition, at least in the cases we rank « good measurements ». Now an elementary calculation shows that the average $\overline{U}$ of

$$\exp[-iHt]\, U \exp[iHt]$$

over a very long time has exactly the form (1) when the observable is the *energy* of the system; and it can be proved that for a big quantum mechanical system, in general U does never deviate much from $\overline{U}$, in the sense that the time average of $(U-\overline{U})^2$ is very small.

Althoug the theory of measurements is not quite satisfactory yet, the way to make it so is clear. The property of a macroscopical system we are looking for is essentially an *ergodic* property: the probability distribution of a physical quantity, as determined by the statistical matrix, does not fluctuate much around its time average, which is its equilibrium value. In other words, the wave function is the superposition of other states whose probability distributions evolve in time without any appreciable interference. From this point of view, Ludwig's approach to the theory of measurements is very sound and does not add anything fundamentally new and unexpected to the theory.

All this, however, can *never remove dispersion*; even if we know through which of the two holes the particle has passed, its position cannot be predicted although its probability distribution has changed. The essential result of any physical measurement is to destroy the interference between the different eigenstates of an observables; but, again, none of such states acquires any privileged standing, capable of distinguishing it as the one we actually observe. In mathematical terms, the transition

$$\sum P_\varepsilon U P_\varepsilon \to P_{\varepsilon_0} U P_{\varepsilon_0} \tag{3}$$

is essentially a probabilistic one and has nothing to do with the time evolution of the system. We must go back again to our main question: what is the meaning of the statement that the only possible description of a physical system is a probabilistic one?

3. – Sometimes one tries to find some solace in the following consideration. Even in classical physics uncertainty reigns uncontested, since any information can be formulated solely in terms of probability; moreover, our knowledge

of any classical system is, in general, unstable [5] in the sense that the law of motion lets an initial uncertainty grow with time without any bound. The purpose and the net outcome of any experiment is just to lessen our ignorance, giving us a chance to start over again with more precise data. In the same way the so-called « quantum jumps » are just the most obvious consequence of the fact that what we know is always only probability. Both classical and quantum physics offer us a *stochastic* model of our knowledge, to wit: a law to carry a probability distribution forward in time.

A closer scrutiny, however, shows an essential difference between the two types of theory. In classical mechanics a stochastic process is generated by a well-defined, deterministic law of motion, so that the state of the system is known with certainty at all times if it is known with certainty at one time. The probabilistic description, in other words, is only a *formal* way to follow up the time development of our ignorance. Among all the different mathematical models actually compatible with our present knowledge there is always one, and only one, compatible also with any possible additional data. That model is called, by definition, the right one and is usually assumed to have something to do with « reality », although its construction is entirely operational and does not imply any metaphysical meaning. Nothing like this occurs in quantum theory. Even if the wave function is known now with certainty, this will never insure a deterministic development; no experiment is accurate enough for this. The quantum physicist has given up any hope of constructing even in principle, a unique model of his experience; nay, he avers that it is meaningless to look for one. An irrational barrier stands, he says, between our curiosity and the motivation of certain simple alternatives occurring every day in a physical laboratory, so that we have to be content with what one could call a « taxonomic » view of nature, according to which all we can do is to classify and to number the various observations.

The comparison with biology, where already the empirical classification of the species is only the first step to a deeper understanding in terms of the theory of evolution, shows how great is our ignorance, although it is shrewdly ma ked by the excellent performance of the quantum statistical laws. One cannot help trying to look at the situation in a humbler way: not being able to explain a class of phenomena, we dress our incapacity with some complicated label and put it at the door of a deep « oddity » of nature. The striking aspect of this attitude is that from the assumed permanence and finality of a scientific theory a relevant conclusion is drawn concerning our knowledge and its

---

(5) See, *e.g.*, M. Born: *Kgl. Dan. Vidensk, Selsk. Mat.-Fys. Medd.*, **30** no. 2; (1955), M. Born and D. J. Hooton: *Kgl. Dan. Vidensk, Selsk. Mat.-Fys. Medd.*, **52**, 287 (1956).

functioning—a problem which has very little to do with physics. This is enough, it seems, to make one suspicious.

4. – Aside from its queer claim on the theory of knowledge, it may be interesting to consider the quantum mechanical model in itself from the point of view of scientific methodology.

Take first the fundamental principle of the correspondence between physical observables and hermitian operators. It may be useful to recall, first of all, that the usual semipictorial description of the quantum phenomena in terms of momentum, position, their uncertainties, and other classical concepts, has no deep significance and should not be allowed to confuse the main issue. We know [1] that the measurement of microscopic quantity entails in fact a complicated chain of statistically irreversible interactions from which, at any stage, a probability distribution for any hermitian operator can be obtained. The definite and unique outcome of the experiment is described by saying that the chain is broken off at a more or less arbitrary point; the probability distribution corresponding to the physical quantity which we observe is then replaced by a δ-function if the measurement had an infinite precision, or a suitable distribution otherwise.

But what is a « *physical quantity* »? How can one construct in principle the operator corresponding to a given « observable »? True, usually the macroscopic dichotomy of a laboratory experiment is so clear-cut (the counter has discharged or not) that in practice the Hilbert space of the state function is naturally split into the sum of a few orthogonal subspaces, corresponding to the different outcomes. This, however, in no way solves the problem because *a*) it may be not enough in some particular situations, and *b*) it leaves the main question of principle without a satisfactory answer. Since every material interaction, from the particle down to the brain cells, is supposedly governed by the ordinary Hamiltonian time evolution, the problem entails the determination of a hermitian operator corresponding to every *perception*. The awkwardness itself of the question suggests, indeed, that we are on a wrong track by relying on mathematically beautiful, but physically unsatisfactory, models, and that we should not go far without a safe *operational* basis [3]. Unfortunately, few physicists push quantum mechanics to these extremes in practice; they rely upon good old classical theories to describe macroscopical phenomena. If momentum or creation, etc., operators are used, it is done only in conjunction with a well-defined experimental apparatus whose working is understood in *classical* terms.

---

[6] See on this point also E. SCHRÖDINGER: *Nuovo Cimento*, **1**, 5 (1955); *Nature*, **173**, 442 (1954).

5. – The other point we wish to make concerns the deep uneasiness of the man of the street when he is told that he must be content with recipes giving the probability of one occurrence or another: « But if I see something, it is there, unique and certain, and I want to understand it. I will never be satisfied with probability and a chronological record of what is happening. »

It may be easy, but it is, in fact, insufficient to answer with the usual positivistic argument: of course we are not concerned with the « real » world, which has no meaning whatsoever; all we really want is an agreement between the theoretical predictions and the experimental observation. The man of the street will not be content with such an answer, which shows a deep gap between physics and human life. As it has been often pointed out by E. SCHRÖDINGER (7), quantum mechanics with its peremptory attitude has climbed into an ivory tower of mathematical abstraction, jeopardizing its role in the human culture.

As far as the scientist is concerned, the « uniqueness of the world » has a precise operational meaning which has nothing to share with metaphysics. The outcome of an experiment is accepted and defined unambiguously simply when it is supported by the majority of the observers. By this, of course, one does not mean that anything about which only few people agree is wrong; we only exclude it, for the time being at least, from our data. Think, for instance, of such « unofficial » sciences as telepathy. A little reflection will show that this criterion of « reality », superficial as it may seem, is really at the back of the mind of every physicist.

Vice versa, one would expect any scientific theory to account for the agreement between the observers on the facts it is concerned with, lest it leaves unexplained the main distinguishing feature of its object. A particle has a chance $\frac{1}{2}$ to be detected by a counter A, and a chance $\frac{1}{2}$ to be detected by a counter B. Why does (to wit, what is the theoretical explanation for the fact that) everybody in the laboratory agree that just the counter A, say, has responded? Does not this agreement suggest that until one has provided for a model of a unique and non-probabilistic link between the particle and *each* percipient we miss the core of the matter? Quantum mechanics deals, one could say (8), with a « Gibbs ensemble of independent observers », since it provides us only with the probability of a given experimental result; but in this case the word « observer » must be carefully explained. The quantum mechanical probability distributions refer really to different experiments, not to different observers; in other words, an « observer » is meant to be « the set of physicists who watch a given instrument at the same time ». In a probabilistic doctrine like quantum theory no provision is made to explain for the

(7) E. SCHRÖDINGER: *Brit. J. Phil. Sc.*, **3**, 109, 233 (1952).

(8) On this order of ideas see H. EVERETT, III: *Rev. Mod. Phys.*, **29**, 454 (1957).

obvious fact that they agree. Again, one cannot help feeling some uneasiness if confronted with the task of giving a final and operational foundation to quantum mechanics: the definition « the set of physicists, etc. » is obviously too vague for this purpose [9].

6. – No proof—either experimental or internal—that quantum mechanics is wrong is known to exist. The problem to be faced with a man-made theory, however, is not to disprove it or not, but whether to accept it or not; that is to say, to judge it in the broader framework of human understanding, and not only check analytically its experimental prediction. We have endeavored to show, along the line of thinking of many other physicists, that one may have some misgivings about its adequacy and logical clarity. The fact that time seems not yet ripe to produce a better substitute should not be a good enough reason to assert its finality. Our knowledge (both experimental and theoretical) of a non-equilibrium process (like those which always occur in any measurement) is too scanty to suggest a better solution to the main problem at stake, that is, the theory of measure, and, more in general, the relationship between the macroscopic world and the elementary particles. But it is certainly important to keep in mind the « *probabilistic uneasiness* » and the question it raises.

* * *

It is my pleasure to thank Dr. R. OPPENHEIMER for his hospitality at the Institute for Advanced Study (Princeton, N. J.). I owe interesting remarks to H. BUCHDAHL and J. PLEBANSKI; above all, however, I feel grateful to Dr. E. SCHRÖDINGER, from whom I learned many of the ideas which led to the present discussion.

---

[9] *Note added in proof*: Prof. SCHRÖDINGER aptly pointed out to me that the agreement between the different pictures of the world present in the minds of different individuals is far from being clear in its full implications, nay it constitutes the « greatest and absolutely unexplicable marvel ».

SUPPLEMENTO AL VOLUME XVII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1960
3° Trimestre

Quantum-Mechanical Many-Particle Treatment of Sound Propagation.

D. E. McCUMBER (*)

Laboratoire de Physique Théorique et Hautes Energies, Faculté des Sciences - Orsay

(ricevuto l'11 Maggio 1960)

CONTENTS. — Introduction. – 1. General formalism. – 2. Transport properties and sound propagation. – 3. Integral equations for the $\mathscr{D}$-functions. 4. Special cases of the approximate equations. – 5. Sound propagation in a nearly-perfect gas. – Appendix.

Introduction.

Considerable progress had been made in recent years in the quantum-mechanical treatment of properties of systems consisting of a large number of interacting particles. With respect to those properties of the systems for which the interaction may be treated as a small perturbation, the formal development of the theory may be said to be complete. However, the theory is by no means complete in the analytic description of properties which cannot be treated by assuming a weak perturbative interaction. Insight into the application of mathematical techniques to such phenomena can frequently be gained by studying physical effects, such as sound propagation, many of whose properties are already described by classical methods. By undertaking such a study, one will also gain information relative to the applicability of the classical arguments based, for example, upon the Boltzmann equation.

In the present work we consider the problem of sound propagation. In the first section we summarize the relevant features of a field theoretic many-particle formalism applicable to the sound problem. The specific formal relations which apply in the treatment of transport properties and of sound propa-

(*) National Science Foundation Postdoctoral Fellow. Present address; Institute for Theoretical Physics, University of Copenhagen, Denmark.

gation in particular are presented in the next section. In the third section we discuss exact and approximate integral equations for the density correlation function or sound propagator. In developing approximate equations we take into account the important time-invariance properties which result from the fact that many relevant macroscopic operators commute with the Hamiltonian. In the fourth section we discuss simple forms of the general class of approximate equations previously developed. We indicate that the mechanism for sound propagation can vary from system to system. We further observe that in « perfect » gases sound propagation is related to single-particle damping, a second-order effect. Having obtained by the general approximation technique a second-order Boltzmann-like equation, we solve that equation for the perfect-gas case in a manner which indicates the importance of the time-invariance properties previously discussed. The approximations used in obtaining the solution have immediate application to more general problems such as equilibrium restoration.

1. – General Formalism.

The formalism to be employed in this paper is an extension of the field theoretic formulation recently presented by P. C. MARTIN and J. SCHWINGER [1]. In addition to developing the necessary extension we very briefly review aspects of that formalism in order to establish a consistent notation.

Assuming that the operators $\{H; Q_i, i=1, ..., m\}$ are not explicitly time dependent and that they form a set of commuting observables, we introduce the fundamental trace

$$\text{(1.1)} \qquad \mathrm{Tr}\,(\alpha_i; \beta) \equiv \exp\,[-W(\alpha_i, \beta)] \equiv \mathrm{Tr}\,(\exp\,[-\beta H - \sum_{i=1}^{m} \alpha_i Q_i])$$

which is analogous to the partition function of statistical mechanics. The function (1.1) is related by Laplace-Stieltjes transform to the density of states characterized by the set of eigenvalues $\{E; q_i, i=1, ..., m\}$ of $\{H; Q_i, i=1, ..., m\}$. Similar relations hold for the other functions to be introduced shortly. Although the results may be immediately generalized to a more general set $\{Q_i\}$, we shall confine ourselves in this work to a conservative Hamiltonian H which commutes with the total numer N of particles and shall assume that the two operators H, N make up the commuting group of (1.1). The resulting trace converges in the half-planes

$$\text{(1.2}a\text{)} \qquad \mathrm{Re}\,\beta > 0\,, \qquad \mathrm{Re}\,\alpha > \mathrm{Re}\,\alpha_0\,,$$

---

[1] P. C. MARTIN and J. SCHWINGER: *Phys. Rev.*, **115**, 1342 (1959).

where

$$\alpha_0 \equiv -\beta \lim_{N\to\infty} \inf \left\{\frac{\mathrm{Tr}\,(P_N H)}{N\,\mathrm{Tr}\,(P_N)}\right\}, \tag{1.2b}$$

and P_N is the projection operator appropriate to N-particle states.

Following MARTIN and SCHWINGER, we introduce the Green's functions

$$G_n(1 \ldots n;\, 1' \ldots n') \equiv (-i)^n \langle(\psi_1 \ldots \psi_n \psi^+_{n'} \ldots \psi^+_{1'})_+ \varepsilon\rangle \tag{1.3}$$

describing the behavior of the fields ψ, ψ^+ in the presence of the background implicit in the « expectation value » $\langle\ldots\rangle$. The time-ordering operation, indicated in (1.3) by $(\ldots)_+$, orders the field operators ψ, ψ^+ from right to left in the increasing order of their time co-ordinates. The factor ε is identically $+1$ for B-E fields, while for F-D fields it is that antisymmetric function of the time coordinates which equals $+1$ when the fields occur in the order exhibited in (1.3). The differential (or more generally the integro-differential) equation satisfied by G_n is governed by the equations of motion of the fields ψ, ψ^+ and indirectly by the expectation value $\langle\ldots\rangle$. The fields ψ, ψ^+ satisfy the equal-time commutation (anticommutation) relations

$$\begin{cases} [\psi(\boldsymbol{r}t),\ \psi\ (\boldsymbol{r}'t)]_\mp = [\psi^+(\boldsymbol{r}t),\ \psi^+(\boldsymbol{r}'t)]_\mp = 0\,, \\ [\psi(\boldsymbol{r}t),\ \psi^+(\boldsymbol{r}'t)]_\mp = \delta(\boldsymbol{r}-\boldsymbol{r}')\,, \end{cases} \tag{1.4}$$

in which internal co-ordinates are implicit. If the internal co-ordinates are used to represent several types of B-E particles and several types of F-D particles, the fields associated with one type of statistics commute with those associated with the other type at the same time. Since the results are easily generalized, we shall restrict ourselves in this work to a single type of particle, indicating within an equation the possibly different B-E/F-D results by the upper and lower of the two expressions.

Appropriate to the (H, N) trace (1.1), we define the expectation value of (1.3) such that

$$\langle X\rangle \equiv \exp\left[W(\alpha\beta)\right]\cdot \mathrm{Tr}\,(\exp\left[-\alpha N - \beta H\right] X)\,. \tag{1.5}$$

The importance of the choice (1.5) rests in the periodicity condition appropriate to G_n in the limiting case $\mathrm{Re}\,\beta = 0^+$, $\mathrm{Im}\,\beta \equiv T > 0$ arbitrary. Namely, for any pair of time coordinates t_i, t'_j such that

$$t_i < t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t'_{j-1}, t'_{j+1}, \ldots, t'_n < t'_j < t_i + T\,, \tag{1.6a}$$

one may readily verify by using the cyclic property of the trace (1.5) that

$$(1.6b)\quad G_n(t_1, \ldots, t_i + T, \ldots, t_n; t_1', \ldots, t_j', \ldots, t_n') = \\ = G_n(t_1, \ldots, t_i, \ldots t_n; t_1', \ldots, t_j' - T, \ldots, t_n') = \\ = (\pm \exp[\alpha])\, G_n(t_1, \ldots, t_i, \ldots, t_n; t_1', \ldots, t_j', \ldots, t_n'),$$

where for simplicity we have indicated only the time co-ordinates.

In many applications we are led to consider functions of a single time difference

$$(1.7)\quad g(t_1 - t_2; \xi_1 \xi_2) \equiv (-i) \langle (X_1(t_1 \xi_1)\, X_2(t_2 \xi_2))_+ \varepsilon \rangle$$

formed from some $(+i)^{n-1} G_n$ in which the various time co-ordinates have been set equal to t_1 or t_2 ($\pm 0^+$). In (1.7) the ε is $+1$ for $t_1 > t_2$ and ± 1 for $t_1 < t_2$ depending upon whether the operators X_i contain an even or an odd number of F-D fields ψ, ψ^+. Using the periodicity property (1.6), we may easily demonstrate that the Fourier transform of (1.7) may be written in the form

$$(1.8)\quad g(\omega; \xi_1 \xi_2) = \frac{\exp[\alpha m + \beta\omega]}{\exp[\alpha m + \beta\omega] - (\pm 1)}\, g(\omega + i\varepsilon; \xi_1 \xi_2)_0 - \\ - \frac{(\pm 1)}{\exp[\alpha m + \beta\omega] - (\pm 1)}\, g(\omega - i\varepsilon; \xi_1 \xi_2)_0, \quad \varepsilon \to 0^+,$$

where $(\pm 1) \equiv \varepsilon(t_1 < t_2)$; $g(\omega + i\varepsilon)_0$ is the Fourier transform of the retarded commutator (anticommutator)

$$(1.9)\quad \begin{cases} g(t_1 - t_2; \xi_1 \xi_2)_0 \equiv (-i) \langle [X_1(t_1 \xi_1),\, X_2(t_2 \xi_2)]_{\mp 1} \rangle\, \theta(t_1 - t_2), \\ \theta(\tau) = \begin{cases} +1 \text{ for } \tau > 0 \\ 0 \text{ for } \tau < 0; \end{cases} \end{cases}$$

and m is the integer determined such that

$$(1.10)\quad m X_1 = [X_1, N] \quad \text{or} \quad m X_2 = [N, X_2].$$

Since H is Hermitian, the Fourier transform of the retarded function (1.9) may be written in the form

$$(1.11)\quad g(\omega + i\varepsilon; \xi_1 \xi_2)_0 = \int_{-\infty}^{+\infty} \frac{d\bar{\omega}}{2\pi} \frac{A_g(\bar{\omega}; \xi_1 \xi_2)}{\omega + i\varepsilon - \bar{\omega}}.$$

Moreover, from (1.8) and (1.11) we have the discontinuities

$$(1.12)\qquad D_n(\xi_1\xi_2) \equiv i^{n+1}\left[\frac{\mathrm{d}^n g}{\mathrm{d}t_1^n}\bigg|_{t_1-t_2=0^+} - \frac{\mathrm{d}^n g}{\mathrm{d}t_1^n}\bigg|_{t_1-t_2=0^-}\right] = \int_{-\infty}^{+\infty}\frac{\mathrm{d}\omega}{2\pi}\,\omega^n A(\omega;\,\xi_1\xi_2)\,; \qquad n = 0, 1, 2, \ldots .$$

In the event of $g(12)$ or $g(12)_0$ representing a system response function, the associated dispersion relations follow immediately from (1.11). The expressions (1.12) represent « sum rules ». By expanding $(\omega \pm i\varepsilon - \bar{\omega})^{-1}$ as a (semi-convergent) power series in $\bar{\omega}/\omega$, we may easily relate the latter to the system high-frequency response.

The functions $A_g(\omega;\,\xi_1\xi_2)$ may display special symmetry properties. For example, if $X_1 = X_2^+$ for $(t_1\xi_1) = (t_2\xi_2)$, then $(\exp[\alpha m + \beta\omega] \mp 1)^{-1} A_g(\omega;\,\xi_1\xi_2)$ for ω, α, β real is a non-negative Hermitian matrix with respect to the indices $(\xi_1\xi_2)$. If $D_0(\xi_1\xi_2) \neq 0$, it is sometimes convenient, in the representation for which $A_g(\omega;\,\xi_1\xi_2)$ is diagonal in the eigenvalues ξ, to represent the function (1.11) for ω *real* in the form

$$(1.13)\qquad g(\omega \pm i\varepsilon;\,\xi_1\xi_2)_0 = \delta(\xi_1\xi_2)\,\frac{D_0(\xi_1)}{\omega - \varepsilon(\omega\xi_1) \pm i\left[\varepsilon + \tfrac{1}{2}\Gamma(\omega\xi_1)\right]},$$

where the functions $D_0(\xi_1)$, $\varepsilon(\omega\xi_1)$, $\Gamma(\omega\xi_1)$ are real [2]. The real-ω form (1.13) is sufficient with (1.11) to determine completely the functions $A(\omega;\,\xi_1\xi_2)$; moreover, the latter determines $g(\omega)_0$ for all ω. As a second example we consider operators X_i such that $X_i^+ = X_i$, $i = 1, 2$. In that case for α, β, ω real

$$(1.14a)\qquad (\exp[-\alpha m + \beta\omega] \mp 1)\,A(\omega;\,\xi_1\xi_2) = -A^*(-\omega;\,\xi_1\xi_2)\cdot(\exp[\alpha m + \beta\omega] \mp 1)$$

or, for $m = 0$,

$$(1.14b)\qquad A(\omega;\,\xi_1\xi_2) = -A^*(-\omega;\,\xi_1\xi_2)\,.$$

If $D_0(\xi_1\xi_2) \neq 0$, then we may again introduce a representation of the type (1.13) for ω real. In this case, however, $D_0(\xi_1)$ is pure imaginary; the functions $\varepsilon(\omega\xi)$, $\Gamma(\omega\xi)$ are real; $\varepsilon(\omega;\,\xi) = -\varepsilon(-\omega;\,\xi)$; and $\Gamma(-\omega;\,\xi) = \Gamma(\omega;\,\xi)$. In the event of the conditions $X_1 = X_2^+$, $X_i^+ = X_i$, $m = 0$ holding simultaneously,

---

[2] For the special case $g(12) = G_1(12)$ explicit approximate forms of $\varepsilon(\omega)$, $\Gamma(\omega)$ have been discussed by MARTIN and SCHWINGER, *op. cit.*

$D_0 \equiv 0$ and the representation (1.13) is inapplicable. In the representation for which $A(\omega; \xi_1 \xi_2)$ is diagonal,

$$A(\omega; \xi_1)_- = -A(-\omega; \xi_1) \tag{1.15a}$$

is real and (1.11) becomes

$$g(\omega + i\varepsilon; \xi_1 \xi_2) = \delta(\xi_1 \xi_2) \int_{-\infty}^{\infty} \frac{d\bar{\omega}}{2\pi} \frac{\bar{\omega} A_g(\bar{\omega}; \xi_1)}{(\omega + i\varepsilon)^2 - \bar{\omega}^2} \,. \tag{1.15b}$$

If $D_1(\xi_1) \neq 0$, we may write this last function for *real* ω in the form

$$g(\omega \pm i\varepsilon; \xi_1 \xi_2) = \frac{\delta(\xi_1 \xi_2) D_1(\xi_1)}{[\omega \pm i(\varepsilon + \frac{1}{2} \Gamma(\omega \xi_1))]^2 - [\varepsilon(\omega \xi_1)]^2} \,, \tag{1.15c}$$

where here $D_1(\xi)$, $[\varepsilon(\omega \xi)]^2$, $\Gamma(\omega \xi)$ are real; $\varepsilon(-\omega; \xi) = \varepsilon(\omega; \xi)$; and $\Gamma(-\omega; \xi) = -\Gamma(\omega; \xi)$. As before, the real-$\omega$ form is sufficient with (1.11) to completely determine A_g.

The function $A_g(\omega; \xi_1 \xi_2)$ is uniquely related to the associated auxiliary function $g(\tau)_0$ and its Fourier trasform $g(\omega)_0$. A similar one-to-one correspondence exists between the function A_g and the function $g(\tau)$ of (1.7) only if the weight function $(\exp[\alpha m + \beta \omega] \mp 1)^{-1}$ does not display a pole on the real-ω axis. If the $(+)$ sign obtains, there is no such pole. However, if the $(-)$ sign obtains, the weight function has a pole at $\omega = -m(\alpha/\beta) \equiv m\mu$. In that case $g(\omega)$ is undetermined to the extent of a term of the form $C(\xi_1 \xi_2) \cdot \delta(\omega - m\mu)$. Since α, β are parameters in the present formalism, one may argue that no such terms exist if $m \neq 0$ [3]. However, the case with $m = 0$ has a real physical significance in that it expresses the fact that the physically meaningful class of functions

$$g_c(t_1 - t_2; \xi_1 \xi_2) = g(t_1 - t_2; \xi_1 \xi_2) + C(\xi_1 \xi_2) \tag{1.16}$$

is associated with a single auxiliary function $g(t_1 - t_2; \xi_1 \xi_2)_0$.

The importance of the auxiliary function $g(\omega)_0$ rests in the following theorem relevant to finite-interval convolution integrals appropriate to the application of Green's theorem to functions having a « periodicity » of the type inherent in (1.8).

---

[3] This non-trivial statement is very relevant to *B-E* condensation. Note that it is *not* equivalent to the assertion that such δ-function terms are not frequently good approximations to the actual $g(\omega)$ in the event that $g(\omega)_0$ or $A(\omega)$ is not known with sufficient accuracy in the vicinity of the weight-function pole.

Theorem (4). – Let $g_i(t_1 - t_2)$, $i = 1, \ldots, n$, be functions such that their Fourier transforms have the form (1.8) with the respective auxiliary functions $g_i(\omega)_0$, $i = 1, \ldots, n$. Then the integral

$$(1.17a) \qquad X(t_1, t_2) \equiv \int_{t<}^{t>} \mathrm{d}t_1'' \ldots \mathrm{d}t_{n-1}'' g_1(t_1 - t_1'') g_2(t_1'' - t_2'') \ldots g_n(t_{n-1}'' - t_2)$$

for $t_< < t_1, t_2 < t_> \equiv t_< + T$, $\mathrm{Re}\,\beta = 0^+$, $\mathrm{Im}\,\beta \equiv T > 0$ is analytically continuable to the whole half plane $\mathrm{Re}\,\beta > 0$ and to the whole real t_1, t_2-axis. The analytic continuation is independent of $t_<$, is a function only of the relative time $t_1 - t_2$, and has a Fourier transform of the form (1.8) with the auxiliary function

$$(1.17b) \qquad . \; X(\omega \pm i\varepsilon)_0 = \prod_{i=1}^{n} g_i(\omega \pm i\varepsilon)_0 \, .$$

It is important to note that the α, β weight functions in (1.8) must be the same for all of the $g_i(\omega)$ in (1.17). In the event of this being not the case or in the event of one of the g_i being replaced by a function such as G_n, $n > 1$, of several time variables, the lemma does not apply, although the result may still have relatively simple periodicity properties. These special cases must be separately treated, generally a tedious but straightforward procedure. Incidentally, note that the usual convolution-integral theorem associated with retarded vacuum functions is the special case of (1.17) for which $\alpha, \beta \to + \infty$.

2. – Transport properties and sound propagation.

Having outlined the general formalism in the preceding section, we turn to the question of transport properties and, specifically, sound propagation. Let $\{O_i(t)\}$ be a set of possibly time-dependent Hermitian observables constructed from the implicitly time-dependent fields ψ, ψ^+. Let the set of real scalar functions $\{F_i(t)\}$ be external time-dependent parameters conjugate to the $\{O_i(t)\}$ and representing perturbating external fields or the response of measurement devices. The Hermitian modification in the Hamiltonian H

(4) Cf. D. E. McCumber: Thesis (Harvard University, 1959); also, P. C. Martin and J. Schwinger: *op. cit.*

genrated by the F_i is, by definition,

$$H(t) = \sum_i F_i(t)\, O_i(t)\,. \tag{2.1}$$

In order for the Laplace-Stiltjes connection between the α, β functions and the eigenstates of N, H to have meaning in the presence of the $\{F_i(t)\}$, it is necessary that ΔH affects the time development of the fields ψ but that it does not modify H in the weighting exponential of the trace (1.5). In essence, the states are to be weighted in (1.5) according to the energy they displayed some time in the distant past ([5]). The time behavior of the $\{F_i\}$ perturbed expectation values

$$\langle O_i(t)\rangle_{\{F\}}\,, \qquad \langle O_i(_1)\, O_j(t_2)\rangle_{\{F\}}\,, \qquad \text{etc..} \tag{2.2}$$

constitues the « transport properties » of the system. For many applications it is sufficient to consider only the first-order modifications induced in $\langle O_i(t_1)\rangle$ by a unit impulse in F_j at time t_2 ([6]). It is easy to verify that that function is

$$\frac{\delta\langle O_i\rangle}{\delta F_j}\,(t_1, t_2) = (-i)\langle [O_i(t_1),\, O_j(t_2)]_-\rangle\,\theta(t_1 - t_2)\,, \tag{2.3}$$

which is a function of the type (1.7) or, more precisely, of the type (1.9). In this case the parameter m of (1.11) is zero; $\varepsilon(t_1 < t_2) \equiv +1$.

Sound propagation may be generated and detected by coupling external fields to the particle-density operator. It follows immediately from (2.3) that the propagation of density fluctuations or sound is described to the first order by the density correlation function

$$\mathscr{D}(t_1 - t_2;\, \boldsymbol{r}_1, \boldsymbol{r}_2)_0 \equiv (-i)\langle [\psi^+\psi(\boldsymbol{r}_1 t_1),\ \psi^+\psi(\boldsymbol{r}_2 t_2)]_-\rangle \cdot \theta(t_1 - t_2) \tag{2.4}$$

or any of the corresponding class of functions (1.16).

In this work we shall restrict our considerations to the bulk properties of homogeneous systems of arbitrarily large volume V. Although in the limit

---

([5]) The question of approximating the expectation values so obtained by those obtained from new ensembles which are characterized by a time-varying Hamiltonian and by local time-dependent variations in the statistical parameters is related to the question of transport of the physical quantities conjugate to those parameters but will not be discussed here.

([6]) For a more complete discussion of the higher-order terms cf. W. Bernard and H. B. Callen: *Rev. Mod. Phys.*, **31**, 1017 (1959).

$V \to \infty$ such bulk properties do not depend upon the boundary conditions, we assume for simplicity that the multiparticle system is confined to a rectangular box on the surface of which periodic boundary conditions obtain. For definiteness we assume that the (non-perturbed) system is governed by the number-conserving Hamiltonian ($\hbar \equiv 1$)

$$H = \sum_{k_1} \varepsilon^0_{k_1} \psi^+_{k_1} \psi_{k_1}(t) + \tfrac{1}{2} \sum_{k_1} \cdots \sum_{k'_2} \psi^+_{k_1} \psi^+_{k_2} v(k_1 k_2; k'_1 k'_2) \psi_{k'_2} \psi_{k'_1} \tag{2.5a}$$

appropriate to two-particle interactions. Translational invariance requires that v be of the form

$$v(k^*_1 k_2; k'_1 k'_2) = v(k_1 k_2; k'_1 k'_2) \cdot \delta(k_1 + k_2; k'_1 + k'_2) \,. \tag{2.5b}$$

We shall further assume that v is *B-E/F-D* symmetrized. In writing (2.4), (2.5) we have suppressed internal co-ordinate indices.

Since the system is translationally invariant, it is convenient to introduce the spatial Fourier transform of the $\mathscr{D}$-functions:

$$\begin{aligned} \mathscr{D}(t_1 - t_2; k) &= V \int \mathrm{d}(r_1 - r_2) \exp\left[-ik\cdot(r_1 - r_2)\right] \mathscr{D}(t_1 - t_2; r_1 - r_2) = \\ &= (-i) \int \frac{\mathrm{d}(r_1 - r_2)}{V} \exp\left[-ik\cdot(r_1 - r_2)\right] \cdot \sum_{k_1} \cdots \sum_{k_4} \delta(k_1 + k_3; k_2 + k_4) \cdot \\ &\cdot \exp\left[ir_1\cdot(k_2 - k_1)\right] \exp\left[ir_2\cdot(k_4 - k_3)\right] \langle (\psi^+_{k_1} \psi_{k_2}(t_1) \psi^+_{k_3} \psi_{k_4}(t_2))_+ \rangle = \\ &= \sum_{k_1} \sum_{k_2} (-i) \langle (\psi^+_{k_1 - k/2} \psi_{k_1 + k/2}(t_1) \psi^+_{k_2 + k/2} \psi_{k_2 - k/2}(t_2))_+ \rangle \end{aligned} \tag{2.6a}$$

$$\equiv \sum_{k_1} \sum_{k_2} \mathscr{D}(t_1 - t_2; k_1 k_2; k) \,. \tag{2.6b}$$

The functions $\mathscr{D}(\tau; r_1 - r_2)$, $\mathscr{D}(\tau; k)$, $\mathscr{D}(\tau; k_1 k_2; k)$ are functions of the type (1.7). The associated spectral density functions $A(\omega)$ are connected by relations similar to (2.6)

$$A(\omega; k) = V \int (\mathrm{d}r) \exp\left[-ik\cdot r\right] A(\omega; r) = \sum_{k_1} \sum_{k_2} A(\omega; k_1 k_2; k) \,. \tag{2.7}$$

Moreover, it is easy to verify that those functions provide examples of the special symmetries discussed in the preceding section. For real ω we have

$$\begin{cases} \omega A(\omega; k_1 k_2; k) = \omega A^*(\omega; k_2 k_1; k) = \\ \qquad\qquad = -\omega A(-\omega; k_2 k_1; -k) \text{ and « non-negative »}, \\ D_0 = [\langle n_{k_1 - k/2} \rangle - \langle n_{k_1 + k/2} \rangle] \delta(k_1 k_2); \end{cases} \tag{2.8a}$$

$$\text{(2.8}b\text{)} \qquad \begin{cases} \omega A(\omega; \boldsymbol{k}) \text{ is a real, non-negative even function of } \omega, \\ D_0 = 0, \; D_1 = \sum_{\boldsymbol{k}_1} \langle n_{\boldsymbol{k}_1} \rangle [\varepsilon^0_{\boldsymbol{k}_1+\boldsymbol{k}} - 2\varepsilon^0_{\boldsymbol{k}_1} + \varepsilon^0_{\boldsymbol{k}_1-\boldsymbol{k}}]; \end{cases}$$

$$\text{(2.8}c\text{)} \qquad A(\omega; \boldsymbol{r}) = V^{-2} \sum_{\boldsymbol{k}} \exp[i\boldsymbol{k}\cdot\boldsymbol{r}] A(\omega; \boldsymbol{k}) \, .$$

As convenient we may utilize representations of the types (1.13) and (1.15).

The familiar types of sound propagation manifest themselves in the function $\mathscr{D}(\omega; \boldsymbol{k})$ by ω-plane singularities which in the limit $\boldsymbol{k} \to 0$ are dominated by one or more «resonances» which become increasingly sharp as $\boldsymbol{k} \to 0$. As shall become apparent in later sections the microscopic mechanism inducing the resonance may differ in various physical situations. In a two-phase system, for example, the sound propagation in each phase may proceed by a different microscopic mechanism.

It is to be noted that the resonances at frequencies different from zero are frequently accompanied by a non-negligible residue at $\omega = 0$. This is important and must be taken into account if one wishes to verify the calculational results by using sum rules, etc. Specifically, consider the case of a dilute M-B gas of weakly interacting particles, a case we shall discuss in some detail later. It is well know that in the limit $k \to 0$ the function $A(\omega; k)$ will exhibit sharply peaked resonances in the vicinity of $\omega = \pm \omega_0$ where

$$\text{(2.9)} \qquad \omega_0^2 = \frac{V\chi_s^{-1}}{m\langle N\rangle} k^2 = \frac{\gamma V \chi_\beta^{-1}}{m\langle N\rangle} k^2 = \gamma k^2 \langle N\rangle \left(\frac{\partial\langle N\rangle}{\partial\mu}\right)^{-1}_{\beta} .$$

In this equation χ_s, χ_β are the isentropic and isothermal compressibilities, λ is the ratio of specific heats, and $\mu = -\alpha/\beta$. We have implicitly assumed in writing (2.9) that $\varepsilon^0_{k_1} = k_1^2/2m$. We also know that

$$\text{(2.10}a\text{)} \qquad \frac{1}{\beta}\frac{\partial\langle N\rangle}{\partial\mu} = \langle N^2\rangle - \langle N\rangle^2 = \lim_{k\to 0} \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} \frac{A(\omega; \boldsymbol{k})}{\exp[\beta\omega] - 1} =$$

$$= \lim_{k\to 0} \frac{1}{\beta} \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{2\pi} \frac{A(\omega; \boldsymbol{k})}{\omega} ,$$

and

$$\text{(2.10}b\text{)} \qquad \lim_{\boldsymbol{k}\to 0} D_1 = \lim_{\boldsymbol{k}\to 0} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega}{2\pi} \omega A(\omega; \boldsymbol{k}) = \lim_{\boldsymbol{k}\to 0} \sum_{\boldsymbol{k}_1} \langle n_{\boldsymbol{k}_1}\rangle \boldsymbol{k}\boldsymbol{k} : \nabla_1 \nabla_1 \varepsilon^0_{\boldsymbol{k}_1} .$$

If we assume that a *single* resonance of the type (2.9) obtains, then we may rewrite (2.10*a*) as

$$\frac{\partial\langle N\rangle}{\partial\mu}(1-c)=\lim_{k\to 0}\frac{1}{\omega_0^2}\int_{-\infty}^{\infty}\frac{\mathrm{d}\omega}{2\pi}\,\omega A(\omega;\boldsymbol{k})\,, \tag{2.10c}$$

where $c(\partial\langle N\rangle/\partial\mu)$ is that contribution to $\lim_{k\to 0}\int(\mathrm{d}\omega/2\pi)(A(\omega;\boldsymbol{k})/\omega)$ which does *not* arise from the $\omega=\pm\omega_0$ resonances. From (2.10*b*) and (2.10*c*) we have immediately for $\varepsilon^0_{k_1}=k_1^2/2m$

$$\omega_0^2=(1-c)^{-1}\left(\frac{\partial\langle N\rangle}{\partial\mu}\right)^{-1}k^2\langle N\rangle\,, \tag{2.11}$$

which corresponds to (2.9) only if $(1-c)^{-1}=\gamma$.

3. – Integrals equations for the $\mathscr{D}$-functions.

The $\mathscr{D}$-functions of the preceding section are special cases of the Green's function iG_2. For example,

$$\mathscr{D}(t_1-t_2;\boldsymbol{k}_1\boldsymbol{k}_2;\boldsymbol{k})\equiv(-i)\langle(\psi^+_{\boldsymbol{k}_{1-}}\psi_{\boldsymbol{k}_{1+}}(t_1)\psi^+_{\boldsymbol{k}_{2+}}\psi_{\boldsymbol{k}_{2-}}(t_2))_+\rangle= \tag{3.1a}$$
$$=iG_2(\boldsymbol{k}_{1+}\,t_1,\ \boldsymbol{k}_{2-}\,t_2;\ \boldsymbol{k}_{1-}\,t_1^+,\ \boldsymbol{k}_{2+}\,t_2^+)\,,$$

where we have introduced (7)

$$\boldsymbol{k}_{j\pm}\equiv\boldsymbol{k}_j\pm\tfrac{1}{2}\boldsymbol{k}\,. \tag{3.1b}$$

The integral equations satisfied by the $\mathscr{D}$-functions thus follow immediately from those satisfied by G_2 (8). Since it is generally impossible to solve those equations exactly, we are primarily concerned here with the development of techniques which will yield sufficiently accurate approximations tot he $\mathscr{D}$-functions.

The G_n for the non-interacting case with $v=0$ in (2.5) are known exactly.

(7) One must carefully distinguish between the $(\pm)$ superscripts and the $(\pm)$ subscripts. The former refer to time ordering and indicate that an infinitesimal positive number is to be added to or subtracted from the time coordinate. The $(\pm)$ subscripts are to be interpreted in the sense (3.1*b*).

(8) The integral equations satisfied by the G_n have been discussed in some detail in reference (1). They are given in one form in (3.8) below.

For that case

$$\mathscr{D}(\omega+i\varepsilon;\ \boldsymbol{k}_1\boldsymbol{k}_2;\ \boldsymbol{k})_0|_{v=0} = \pm\,\delta(\boldsymbol{k}_1\boldsymbol{k}_2)[iG_1^0(\boldsymbol{k}_{1+};\ t_1-t_2)G_1^0(\boldsymbol{k}_{1-};\ t_2-t_1)](\omega+i\varepsilon)_0 = \tag{3.2}$$

$$= \delta(\boldsymbol{k}_1\boldsymbol{k}_2)\,\frac{f(\varepsilon^0_{k_{1-}})-f(\varepsilon^0_{k_{1+}})}{\omega+\varepsilon^0_{k_{1-}}-\varepsilon^0_{k_{1+}}+i\varepsilon}\,,$$

where

$$f(\omega) \equiv (\exp[\alpha+\beta\omega]\mp 1)^{-1}\,. \tag{3.3}$$

Here, as in the remainder of this work, the upper sign refers to the case where ψ, ψ^+ represent *B-E* fields, the lower sign *F-D* fields. One may easily verify that (3.2) does *not* display a sound mode such as is normally associated with a « perfect » gas. That this is not in contradiction with experience is clear if we recall that the mean free path in « pefect » gases may be macroscopically exceedingly short. The gas represented by (3.2) has an infinitely long mean free path and is therefore a limiting case of a (non-sound-propagating) Knudsen gas. The simplest account of $v \neq 0$ follows by approximating G_2 by a symmetrized product of real G_1's. One may readily demonstrate that such an approximation is also not sufficiently accurate to describe sound propagation. Moreover, if the G_1's include damping effects, the resulting approximation will be strongly damped even in the limit $\boldsymbol{k} \to 0$.

This latter observation is particularly noteworthy since it leads to a significant group of restrictions wich must be imposed upon approximations to the $\mathscr{D}$-functions [9]. For the case $\boldsymbol{k} = 0$ we have from (2.6) and (3.1)

$$\begin{cases} \mathscr{D}(t_1-t_2;\ \boldsymbol{k}_1\boldsymbol{k}_2;\ 0) = (-i)\langle(n_{\boldsymbol{k}_1}(t_1)\,n_{\boldsymbol{k}_2}(t_2))_+\rangle\,; \\ \sum\limits_{\boldsymbol{k}_1}\mathscr{D}(t_1-t_2;\ \boldsymbol{k}_1\boldsymbol{k}_2;\ 0) = (-i)\langle(N(t_1)\,n_{\boldsymbol{k}_2}(t_2))_+\rangle\,; \\ \mathscr{D}(t_1-t_2;\ 0) = (-i)\langle(N(t_1)\,N(t_2))_+\rangle\,. \end{cases} \tag{3.4}$$

Since $[N, H] = 0$, it is clear that the last two functions must be time independent. Similarly, $[\boldsymbol{P}, H] = 0$ implies that

$$\sum_{\boldsymbol{k}_1}\boldsymbol{k}_1\mathscr{D}(t_1-t_2;\ \boldsymbol{k}_1\boldsymbol{k}_2;\ 0) = (-i)\langle(P(t_1)\,n_{\boldsymbol{k}_2}(t_2))_+\rangle \tag{3.5}$$

[9] The remarks which follow are related to the fact that in a multiparticle system the transport cross section differs from the scattering cross section in the role played by the forward scatterings. Cf. the following treatments of electrical conductivity: W. Kohn and J. Luttinger: *Phys. Rev.*, **108**, 590 (1957); S. F. Edwards: *Phil. Mag.*, **3**, 1020 (1958); A. A. Abrikosov and L. P. Gor'kov: *Žurn. Eksp. Teor. Fiz.*, **35**, 1558 (1958) [translation: *Journ. Exp. Theor. Phys.* (*USSR*), **8**, 1090 (1959)].

be time independent. Additional requirements are imposed by the other invariants. Physically it is clear that those invariant attributes which are most important are those which, if not the sum of strictly local operators, are the sum of operators which are local relative to the wavelength k^{-1}. Thus, we must include the restriction imposed by the trivial commutator $[H, H]=0$. For an N-particle system we would have many invariants, a large number of which would not be excluded by the requirement of localizability. It is obvious, of course, that we cannot hope to treat all such invariants exactly; therefore, we must content ourselves with taking only a few of the most important into account. This procedure is perfectly compatible with the statistical mechanical basis for the introduction of α, β functions. In fact, it is clear that the operators which are most important for the present considerations are precisely those operators $\{H; Q_i, i = 1, \ldots, r\}$ which one might be led to introduce into the fundamental trace (1.1). In addition to N, $\boldsymbol{P}$, H, the latter might include the total charge Q, the total number N_i of particles of the i-th type in a non-reacting mixture, the total spin $\boldsymbol{S}$, and the total angular momentum $\boldsymbol{J}$.

To develop systematic approximations to $\mathscr{Q}$ or G_2 in which the postulated invariance properties obtain, we shall assume that the interaction v may be expressed as a function of a parameter ν such that the power series

$$v(12; 1'2')_\nu = \sum_{m=1}^{\infty} \nu^m w_m(12; 1'2') \tag{3.6}$$

and

$$G_n(1 \ldots n; 1' \ldots n')_\nu = \sum_{m=0}^{\infty} \nu^m G_n^m(1 \ldots n; 1' \ldots n') \tag{3.7a}$$

converge in a « sufficiently large » neighborhood of $\nu = 0$. For $\nu = 0$ the functions G_n reduce to the $v = 0$ functions

$$G_n^0(1 \ldots n; 1' \ldots n') = \begin{Bmatrix} \text{perm} \\ \det \end{Bmatrix} [G_1^0(ij')]. \tag{3.7b}$$

The functions w_m in (3.6) are Hermitian, are B-E/F-D symmetrized, and satisfy the same conservation laws (cfr. (2.5b)) as v [10].

[10] The generalization to the interaction (3.6) is not purely academic since such an expansion is, for example, the basis of Brueckner's t-matrix treatment of singular interactions.

The functions G_n satisfy the integral equations (exact)

$$(3.8a)\qquad G_n(1\dots n;\,1'\dots n') = G_1^0(11'')\,G_{n-1}(2\dots n;\,2''\dots n'')\,\frac{\delta_{(n)}(1''\dots n'';\,1'\dots n')}{(n-1)!} \pm$$

$$\pm\, iG_1^0(11'')\,v(1''n+1;\,1'''n+1'')\,G_{n+1}(1'''2\dots nn+1'';\,1'\dots n'n+1^+)\,,$$

$$(3.8b)\qquad = \frac{\delta_{(n)}(1\dots n;\,1''\dots n'')}{(n-1)!}\,G_1^0(1''1')\,G_{n-1}(2''\dots n'';\,2'\dots n') \pm$$

$$\pm\, G_{n+1}(1\dots nn+1^-;\,1''2'\dots n'n+1'')\,v(1''n+1'';\,1'''n+1)\,G_1^0(1'''1')\,.$$

In writing these equations we have used an abbreviated notation; summation or integration over intermediate variables is implicit [11]. Sustituting the series (3.6) and (3.7) into (3.8) we have upon equating coefficients of equal powers of v that for $m \geqslant 1$

$$(3.9a)\qquad G_n^m(1\dots n;\,1'\dots n') = G_1^0(11'')\,G_{n-1}^m(2\dots n;\,2''\dots n'')\,\frac{\delta_{(n)}(1''\dots n'';\,1'\dots n')}{(n-1)!} \pm$$

$$\pm \sum_{r=1}^{m} G_1^0(11'')\,w_r(1''n+1;\,1'''n+1'')\,G_{n+1}^{m-r}(1'''2\dots nn+1'';\,1'\dots n'n+1^+) =$$

$$(3.9b)\qquad = \frac{\delta_{(n)}(1\dots n;\,1''\dots n'')}{(n-1)!}\,G_1^0(1''1')\,G_{n-1}^m(2''\dots n'';\,2'\dots n') \pm$$

$$\pm i \sum_{r=1}^{m} G_{n+1}^{m-r}(1\dots nn+1^-;\,1''2'\dots n'n+1'')\,w_r(1''n+1'';\,1'''n+1)\,G_1^0(1'''1')\,.$$

These equations together with (3.7b) are sufficient to completely determine the G_n^m.

Let us consider in more detail the invariance properties previously discussed. In what follows we mention only G_2 since that function is of immediate interest. It is nevertheless clear that our remarks may be easily extended to the general G_n. With respect to $[N, H] = 0$, we note that

$$(3.10a)\qquad \langle (N(t_1)\,n_{\boldsymbol{k}_2}(t_2))_+ \rangle = \sum_{\boldsymbol{k}_1} \langle (n_{\boldsymbol{k}_2}(t_1)\,n_{\boldsymbol{k}_1}(t_2))_+ \rangle = \sum_{\boldsymbol{k}_1} i^2 G_2(\boldsymbol{k}_1 t_1,\,\boldsymbol{k}_2 t_2;\,\boldsymbol{k}_1 t_1^+,\,\boldsymbol{k}_2 t_2^+)\,.$$

Similary, with respect to $[\boldsymbol{P}, H] = 0$ we have

$$(3.10b)\qquad \langle (\boldsymbol{P}(t_1)\,n_{\boldsymbol{k}_2}(t_2))_+ \rangle = \sum_{\boldsymbol{k}_1} \boldsymbol{k}_1 i^2 G_2(\boldsymbol{k}_1 t_1,\,\boldsymbol{k}_2 t_2;\,\boldsymbol{k}_1 t_1^+,\,\boldsymbol{k}_2 t_2^+)\,.$$

[11] Time integrations extend over the finite interval $(t_<, t_>)$ mentioned in conjunction with (1.17).

The expressions (3.10) have the general form

$$\sum_{\boldsymbol{k}_1} \left\{ \{ \mathscr{A}(\boldsymbol{k}_1; \bar{t}_1 \bar{\bar{t}}_1) \cdot i^2 G_2(\boldsymbol{k}_1 \bar{t}_1, \boldsymbol{k}_2 t_2; \boldsymbol{k}_1 \bar{\bar{t}}_1^{+}, \boldsymbol{k}_2 t_2^{+}) \}_{\bar{t}_1 = \bar{\bar{t}}_1 = t_1} + g(\boldsymbol{k}_1 \boldsymbol{k}_2) \right\}, \tag{3.11}$$

where $\mathscr{A}(\boldsymbol{k}_1; \bar{t}_1 \bar{\bar{t}}_1)$ is a ν-independent function of $(\boldsymbol{k}_1; \bar{t}_1 \bar{\bar{t}}_1)$ and possibly a finite number of their derivatives. The treatment of $[H, H] = 0$ may be reduced to the consideration of a function of the type (3.11) if we observe that

$$\sum_{\boldsymbol{k}_1} \left(2\varepsilon^0_{\boldsymbol{k}_1} + i \frac{\partial}{\partial \bar{\bar{t}}_1} - i \frac{\partial}{\partial t_1} \right) \psi^{+}_{\boldsymbol{k}_1}(\bar{\bar{t}}_1) \psi_{\boldsymbol{k}_1}(\bar{t}_1) \,|_{\bar{t}_1 = \bar{\bar{t}}_1 = t_1} = \tag{3.12}$$

$$= \sum_{\boldsymbol{k}_1} 4\varepsilon^0_{\boldsymbol{k}_1} \psi^{+}_{\boldsymbol{k}_1} \psi_{\boldsymbol{k}_1}(t_1) + 2 \sum_{\boldsymbol{k}_3} \cdots \sum_{\boldsymbol{k}_6} \psi^{+}_{\boldsymbol{k}_3} \psi^{+}_{\boldsymbol{k}_4} v(\boldsymbol{k}_3 \boldsymbol{k}_4; \boldsymbol{k}_5 \boldsymbol{k}_6) \psi_{\boldsymbol{k}_6} \psi_{\boldsymbol{k}_5}(t_1) = 4H(t_1) .$$

Thus,

$$\langle (H(t_1) n_{\boldsymbol{k}_2}(t_2))_{+} \rangle = \tag{3.13}$$

$$= \frac{1}{4} \sum_{\boldsymbol{k}_1} \left\{ \left| \left(2\varepsilon^0_{\boldsymbol{k}_1} + i \frac{\partial}{\partial \bar{\bar{t}}_1} - i \frac{\partial}{\partial t_1} \right) i^2 G_2(\boldsymbol{k}_1 \bar{t}_1, \boldsymbol{k}_2 t_2; \boldsymbol{k}_1 \bar{\bar{t}}_1^{+}, \boldsymbol{k}_2 t_2^{+}) \right|_{\bar{t}_1 = \bar{\bar{t}}_1 = t_1} - \right.$$

$$\left. - 2i \langle n_{\boldsymbol{k}_2} \rangle (1 + \delta(\boldsymbol{k}_1 \boldsymbol{k}_2)) \right\},$$

which is of the asserted form. To proceed, we shall assume that the treatment of all relevant commutators $[Q_i, H] = 0$ may be reduced to the consideration of a function of the type (3.11). We remark that we have here considered only the indices (11′). Since corresponding results obtain for the indices (22′), it is clear that the indices (11′) and (22′) must be treated symmetrically.

Since the function $\mathscr{A}(\boldsymbol{k}_1; \bar{t}_1 \bar{\bar{t}}_1)$ of (3.11) is ν-independent and since for all ν the exact G_2 has the postulated invariance properties, it follows that for each $m \geqslant 0$ the function $G_2^m(12; 1'2')$ given by (3.9) identically satisfies the invariance requirements. Hence, if we require each approximation to G_2 to satisfy those requirements exactly, it is *sufficient* to take the truncated series

$$G_2(12; 1'2'; M)_\nu = \sum_{m=0}^{M} \nu^m G_2^m(12; 1'2') \tag{3.14}$$

as the order-M approximation to G_2. Neglecting for the moment the (11′), (22′) symmetry, we observe that with respect to the indices (11′) the invariance properties will remain if each of the

$$G_2^m(\boldsymbol{k}_1 t_1, \boldsymbol{k}_2 t_2; \boldsymbol{k}_1' t_1', \boldsymbol{k}_2' t_2') \tag{3.15a}$$

in (3.14) is replaced by

$$(3.15b) \qquad \sum_{k_2''} \sum_{k_2'''} \int_{t<}^{t>} \mathrm{d}t_2'' G_2^m(k_1 t_1,\ k_2'' t_2'';\ k_1' t_1',\ k_2''' t_2''^{+}) \cdot X_M^m(k_2''' t_2'',\ k_2 t_2;\ k_2'' t_2''^{+},\ k_1' t_1') \,.$$

This fact follows immediately from the cyclic property of the trace and from the commutators $[Q_j, H] = [Q_i, Q_j] = 0$. Thus, we may replace (3.14) by the alternate approximation

$$(3.16) \qquad G_2(12; 1'1'; M) \equiv \sum_{m=0}^{M} \nu^m \big[G_2^m(12; 1'2') + \\ + G_2^m(1, k_2'' t_2''; 1', k_2''' t_2''^{+}) X_M^m(k_2''' t_2'', 2; k_2'' t_2'', 2') \big],$$

in which we have again utilized a symbolic notation. In order that (3.16) be a systematic approximation to G_2, the functions X_M^m must be such that for each m

$$(3.17) \qquad X_M^m \to 0 \qquad \text{as} \qquad M \to \infty \,.$$

Since X_M^m is possibly ν-dependent, the functions (3.16) are not simply partial sums of a power series expansion in ν. To emphasize this difference from (3.24), we have dropped the suffix ν from the function $G_2(M)$ of (3.16). In addition to having the asymptotic form (3.17), the functions X_M^m must also be such that the symmetry between (11′) and (22′) is maintained. These restrictions still admit a rather large class of functions X_M^m.

In general, the functions (3.16) will not be properly *B-E/F-D* symmetrized, except to the perturbative accuracy in ν of that equation. Moreover, those functions may not be artificially symmetrized without destroying the time invariance required by $[Q_i, H] = 0$. One may show that by a suitable choice of the functions w_n of (3.6) one can often restore all or part of the *B-E/F-D* symmetry. In this regard we should also note that equations (3.14) and (3.16) may be very much more complicated approximations to the real G_2 than a simple perturbative expansion in powers of the interaction v. This is particularly true if the w_n are themselves functionals of G_2. Such considerations are, however, beyond the scope of the present work. We shall here assume that the system under consideration is such that the exchange effects omitted by the improper symmetrization of (3.16) are negligible and that the w_n are independent of the $G_2(M)$.

Depending upon the nature of the w_n in (3.6), the quasi-perturbative results (3.14) will not suffice for a treatment of the collective modes in G_2 appropriate to sound propagation. However, the introduction of a function X

in (3.16) permits us to replace (3.14) by a linear integral equation which, when iterated, will generate the sound modes. In considering the sound modes we are primarily interested in the spectral characteristics of the equal-time function (3.1). This observation suggests that we utilize the function X of (3.16) to shift from a direct expansion such as (3.14) to an expansion for G_2 in which the primary emphasis lies in the spectral behavior. The single-particle Green's function $G_1(11')$ provides a simple clear cut example with which to illustrate such a shift in emphasis. Setting $n = 1$ in (3.8), using theorem (1.17), and recognizing that in the present case G_1 is diagonal in $\boldsymbol{k}$-space, we have

$$(3.18a) \qquad G_1(\boldsymbol{k}_1; \omega + i\varepsilon)_0 = G_1^0(\boldsymbol{k}_1; \omega + i\varepsilon)_0 + \\ + G_1^0(\boldsymbol{k}_1; \omega + i\varepsilon)_0[\pm iv(13; 1''3'')\, G_2(1''3''; 1'3^+)](\boldsymbol{k}_1; \omega + i\varepsilon)_0$$

or, alternatively,

$$(3.18b) \qquad G_1(\boldsymbol{k}_1; \omega + i\varepsilon)_0^{-1} = G_1^0(\boldsymbol{k}_1; \omega + i\varepsilon)_0^{-1} - \\ - [\pm iv(13; 1''3'')\, G_2(1''3''; 1'''3^+)\, G_1(1'''1')^{-1}](\boldsymbol{k}_1; \omega + i\varepsilon)_0 .$$

Expanding v as in (3.6), we are led from (3.18a) to approximate G_1 by the truncated series derived from (3.7)

$$(3.19) \qquad G_1(\boldsymbol{k}; \omega; M) \equiv \sum_{m=0}^{M} v^m\, G_1^m(\boldsymbol{k}; \omega) .$$

This expression is analogous to (3.14). If on the other hand we were to express $[\pm iv G_2\, G_1^{-1}]$ as a power series in v, we would be led to an analogue of the postulated iteration of G_2:

$$(3.20) \qquad G_1(\boldsymbol{k}; \omega + i\varepsilon; M)_0 = G_1^0(\boldsymbol{k}; \omega + i\varepsilon)_0 + \\ + \Big\{\sum_{m=1}^{M} v^m C_1^m(\boldsymbol{k}; \omega + i\varepsilon)_0\Big\} \cdot G_1(\boldsymbol{k}; \omega + i\varepsilon; M)_0,$$

where the coefficients C_1^m of the expansion

$$(3.21) \qquad [\pm G_1^0 iv G_2 G_1^{-1}](\boldsymbol{k}; \omega + i\varepsilon)_0 \equiv \sum_{m=0}^{\infty} v^m\, C_1^m(\boldsymbol{k}; \omega + i\varepsilon)_0$$

may be determined by substituting (3.7) and (3.21) into (3.18) and by equating coefficients of equal powers of v. As one may easily verify,

$$(3.22) \qquad C_1^0 = 0; \qquad C_1^1 G_1^0 = G_1^1; \qquad C_1^M G_1^0 = G_1^m - \sum_{l=1}^{m-1} C_1 G_1^{m-l}, \qquad m \geqslant 2 .$$

Note that C_1^M depends only upon the G_1^m, $m = 0, \ldots, M$.

Suppressing for the moment a detailed consideration of indices, we may easily write the corresponding relations for G_2. Corresponding to (3.18) there is the equation

$$G_2 = G_2^0 + [G_2 - G_2^0]\, G_2^{-1} \cdot G_2 \,. \tag{3.23}$$

If we define coefficients C_2^m such that

$$[G_2 - G_2^0]\, G_2^{-1} \equiv \sum_{m=0}^{\infty} \nu^m\, C_2^m \,, \tag{3.24a}$$

then as in (3.22)

$$C_2^0 = 0; \qquad C_2^1 G_2^0 = G_2^1; \qquad C_2^m = G_2^m - \sum_{l=1}^{m-1} C_2^l\, G_2^{m-l} \,, \qquad m \geqslant 2 \,. \tag{3.24b}$$

Using these results in (3.23), we obtain the order-M approximation analogous to (3.20):

$$G_2(M) = G_2^0 + \Big\{ \sum_{m=1}^{M} \nu^m\, C_2^n\, G_2^0 \Big\}\, (G_2^0)^{-1} \cdot G_2(M) = \tag{3.25a}$$

$$= G_2^0 + \big\{ \nu G_2^1 + \nu^2 [G_2^2 - G_2^1 (G_2^0)^{-1} G_2^1] + \ldots + \nu^M [\ldots] \big\}\, (G_2^0)^{-1}\, G_2(M) \,. \tag{3.25b}$$

This equation is easily seen to be of the same general form as (3.16) provided that the intermediate indices are treated in the equal-time manner of (3.15).

It remains to consider the intermediate indices of (3.25) in some detail since there is some difficulty in defining an inverse $(G_2^0)^{-1}$ which operates solely upon the indices $(11')$ or $(22')$ of $G_2^0(12;\, 1'2')$. Recall from (3.7b) that for our present system

$$G_2^0(12;\, 1'2') = G_1^0(\boldsymbol{k}_1 \xi_1;\, t_1 - t_1')\, G_1^0(\boldsymbol{k}_2 \xi_2;\, t_2 - t_2')\, \delta(\boldsymbol{k}_1 \boldsymbol{k}_1')\, \delta(\boldsymbol{k}_2 \boldsymbol{k}_2')\, \delta(\xi_1 \xi_1')\, \delta(\xi_2 \xi_2') \pm \tag{3.26}$$

$$\pm\, G_1^0(\boldsymbol{k}_1 \xi_1;\, t_1 - t_2')\, G_1^0(\boldsymbol{k}_2 \xi_2;\, t_2 - t_1')\, \delta(\boldsymbol{k}_2 \boldsymbol{k}_1')\, \delta(\boldsymbol{k}_1 \boldsymbol{k}_2')\, \delta(\xi_1 \xi_2')\, \delta(\xi_2 \xi_1') \,,$$

where we have explicitly indicated internal coordinates ξ. As is clear from (3.1), we are primarily interested in the case

$$\boldsymbol{k}_1 - \boldsymbol{k}_1' = \boldsymbol{k}_2' - \boldsymbol{k}_2 \equiv \boldsymbol{k} \neq 0 \,, \tag{3.27}$$

which for simplicity we shall simply designate in what follows as $(\boldsymbol{k} \neq 0)$. For that case (3.26) becomes

$$G_2^0(12;\, 1'2') = \pm G_1^0(\boldsymbol{k}_1 \xi_1;\, t_1 - t_2')\, G_1^0(\boldsymbol{k}_2 \xi_2;\, t_2 - t_1')\, \delta(\boldsymbol{k}_1 \boldsymbol{k}_2')\, \delta(\boldsymbol{k}_2 \boldsymbol{k}_1')\, \delta(\xi_1 \xi_2')\, \delta(\xi_2 \xi_1') \tag{3.28a}$$

and

$$(3.28b) \qquad \delta(12')\,\delta(21') = \pm\, G_1^0(11'')^{-1}\, G_2^0(1''2;\, 1'''2')\, G_1^0(1'''1')^{-1} = \\ = \pm\, G_1^0(22'')^{-1}\, G_2^0(12'';\, 1'2''')\, G_1^0(2'''2')^{-1}\,.$$

That is, for $(\boldsymbol{k} \neq 0)$ we are able to define a simple inverse operator which acts only upon the indices $(11')$ or $(22')$ of $G_2^0(12;\, 1'2')$. We shall therefore assume in the equations that follow that $(\boldsymbol{k} \neq 0)$. Conservation of total momentum insures that an equivalent inequality holds at each intermediate stage where $(G_2^0)^{-1}$ acts. We shall indicate shortly that the $(\boldsymbol{k} \neq 0)$ requirement is not a serious restriction.

As has been noted, the preservation of the time-invariance properties requires that the intermediate indices be treated in the equal-time manner (3.15) ([12]). Moreover, we must meet the requirement (3.27). If we symbolically indicate intermediate stages of the type (3.15) by $*$, the following is an admissible symmetric equation of the type (3.25):

$$(3.29) \qquad G_2(M)\big|_{k\neq 0} = G_2^0 + \left\{\nu G_2^1 + \nu^2[G_2^2 - G_2^1 * (G_2^0)^{-1} * G_2^1] + \ldots + \nu^M[G_2^M - \ldots]\right\} + \\ + \left\{\nu G_2^1 + \ldots\right\} * (G_2^0)^{-1} * G_2(M) * (G_2^0)^{-1}\left\{\nu G_2^1 + \ldots\right\}.$$

The $*$-notation provides a convenient distinction between equal-time operations of the type (3.15) and different ones which might arise in subsequent applications of (3.29). It is easily seen that, at least in so far as the expansion (3.7) has meaning, $G_2(M)_{k\neq 0}$ approaches $G_2(\boldsymbol{k}\neq 0)$ as $M \to \infty$, the error in $G_2(M)$ being perturbatively of order ν^{M+1}. Of course, by virtue of the infinite order iterations implicit in (3.29) the approximate results based upon that equation may contain collective effects not present in the results of the same perturbative accurancy obtained from (3.14).

Since in order to define the inverse $(G_2^0)^{-1}$ we restricted our considerations to the case $(\boldsymbol{k} \neq 0)$, it might appear that equation (3.29) has no value in the case for which $(\boldsymbol{k} = 0)$. In fact, that equation does provide an approximation to that component of $G_2(\boldsymbol{k}=0)$ which varies « continuously » as $\boldsymbol{k} \to 0$ ([13]). For the exact G_2 that component is precisely $G_2(12;\, 1'2') - G_1(11')G_1(22')$.

([12]) Strictly, the expressions (3.25) would have the form (3.16) if the iterative process were defined in a special manner such that only the first and last intermediate integrations were of the type (3.15).

([13]) In figuratively mentioning continuity we of course neglect the discrete nature of the momenta, although that may be extremely important (as in low temperature B-E systems). The distinction to be made is more precisely between those terms of G_2 which contain for all α, β the factor $\delta(\boldsymbol{k}_1, \boldsymbol{k}_1')\,\delta(\boldsymbol{k}_2, \boldsymbol{k}_2')$ and those which do not. The « limit » of (3.30) is to be understood in the same sense.

Thus we are led to define more generally [13]

$$(3.30)\qquad G_2(12;\,1'2';\,M) = \left[G_1(11')\,G_1(22')\right]_M + \lim_{k''\to k} G_2(12;\,1'2';\,M)_{k''\neq 0}$$

as the order-M approximation valid for arbitrary momenta. In this equation $G_2(M)_{k''\neq 0}$ is to be determined from (3.29); $(G_1G_1)_M$ represent an order-M approximation to G_1G_1. It might be computed from (3.20) or, alternatively, it might be computed from the « discontinuous » component of $G_2(12;\,1'2')_{k\neq 0}$ associated with the δ-functions $\delta(\boldsymbol{k}_1\,\boldsymbol{k}_2')\,\delta(\boldsymbol{k}_2\,\boldsymbol{k}_1')$.

4. – Special cases of the approximate equations.

In what follows we shall restrict ourselves to the equal-time ($t_1 = t_1'$; $t_2 = t_2'$) form of equation (3.29). Each of the components of such an equation is a function of the type (1.7) having the same periodicity as the $\mathscr{D}$-functions of Section 2. It follows that the finite-interval Fourier integral theorem (1.17) applies to the time coordinates.

Corresponding to the $G_2(M)$ defined by (3.29), we introduce the order-M approximation

$$(4.1)\qquad \mathscr{D}_M(t_1 - t_2;\,\boldsymbol{k}_1\boldsymbol{k}_2;\,\boldsymbol{k}) \equiv iG_2(\boldsymbol{k}_{1+}\,t_1,\,\boldsymbol{k}_{2-}\,t_2;\,\boldsymbol{k}_{1-}\,t_1^+,\,\boldsymbol{k}_{2+}\,t_2^+;\,M)$$

to the $\mathscr{D}$-function (3.1). It is convenient to introduce the slightly more general function

$$(4.2)\qquad \mathscr{D}_M(t_1 - t_2;\,\boldsymbol{k}_1\,\boldsymbol{k}_1',\,\boldsymbol{k}_2\,\boldsymbol{k}_2') \equiv iG_2(\boldsymbol{k}_1 t_1,\,\boldsymbol{k}_2 t_2;\,\boldsymbol{k}_1' t_1^+,\,\boldsymbol{k}_2' t_2^+;\,M)$$

and the components

$$(4.3)\qquad \mathscr{D}^m(t_1 - t_2;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2') \equiv iG_2^m(\boldsymbol{k}_1 t_1,\,\boldsymbol{k}_2 t_2;\,\boldsymbol{k}_1' t_1^+,\,\boldsymbol{k}_2' t_2^+),$$

which would appear in an expansion of the type (3.7). Using these definitions with theorem (1.17) in equation (3.29), we have [14]

$$(4.4a)\qquad \mathscr{D}_M(\omega + i\varepsilon;\,\boldsymbol{k}_1\,\boldsymbol{k}_1',\,\boldsymbol{k}_2\,\boldsymbol{k}_2')_0\,|_{\boldsymbol{k}\neq 0} =$$

$$= \mathscr{D}^0(\omega + i\varepsilon;\,\boldsymbol{k}_1\,\boldsymbol{k}_1',\,\boldsymbol{k}_2\,\boldsymbol{k}_2')_0 + \sum_{\boldsymbol{k}_2''} \cdots \sum_{\boldsymbol{k}_2^{v}} \{\ldots\}\,(\omega + i\varepsilon;\,\boldsymbol{k}_1\,\boldsymbol{k}_1',\,\boldsymbol{k}_2''\,\boldsymbol{k}_2''')_0 \cdot$$

$$\cdot\,\mathscr{D}^0(\omega + i\varepsilon;\,\boldsymbol{k}_2'''\,\boldsymbol{k}_2'',\,\boldsymbol{k}_2^{iv}\,\boldsymbol{k}_2^{v})_0^{-1} \cdot \mathscr{D}_M(\omega + i\varepsilon;\,\boldsymbol{k}_2^{v}\,\boldsymbol{k}_2^{iv},\,\boldsymbol{k}_2\,\boldsymbol{k}_2')_0 =$$

[14] In dealing with the auxiliary function $\mathscr{D}(\omega + i\varepsilon)_0$ we no longer need to indicate $(\boldsymbol{k}\neq 0)$ since the $(\boldsymbol{k}=0)$ and the $(\boldsymbol{k}\neq 0)$ $\mathscr{D}$-functions both belong to the same class (1.16). However, it is sometimes convenient to indicate the $(\boldsymbol{k}\neq 0)$ explicitly as a reminder that no δ-function component should be added in computing $\mathscr{D}(\omega)$ for $(\boldsymbol{k}\neq 0)$ from $\mathscr{D}(\omega \pm i\varepsilon)_0$.

$$(4.4b)\qquad -\mathscr{D}_{M-1}(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2')_0+\sum_{\boldsymbol{k}_2''}\cdots\sum_{\boldsymbol{k}_2^{v}}v^M\big[\mathscr{D}^M(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2''\boldsymbol{k}_2''')_0-\ldots\big]$$
$$\cdot\mathscr{D}^0(\omega+i\varepsilon;\,\boldsymbol{k}_2'''\boldsymbol{k}_2'',\,\boldsymbol{k}_2'^{v}\boldsymbol{k}^{v})_0^{-1}\cdot\mathscr{D}_M(\omega+i\varepsilon;\,\boldsymbol{k}_2^{v}\boldsymbol{k}_2'^{v},\,\boldsymbol{k}_2\boldsymbol{k}_2')_0,$$

where

$$(4.4c)\qquad \{\ldots\}(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2')_0\equiv$$
$$v\mathscr{D}'(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2')_0+v^2\big[\mathscr{D}^2(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2')_0-\sum_{k_2}\cdots\sum_{k_2^{v}}\cdot$$
$$\cdot\mathscr{D}^1(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2''\boldsymbol{k}_2''')\cdot\mathscr{D}^0(\omega+i\varepsilon;\,\boldsymbol{k}_2'''\boldsymbol{k}_2'',\,\boldsymbol{k}_2'^{v}\boldsymbol{k}_2^{v})_0^{-1}\cdot$$
$$\cdot\mathscr{D}^1(\omega+i\varepsilon;\,\boldsymbol{k}_2^{v}\boldsymbol{k}_2'^{v},\,\boldsymbol{k}_2\boldsymbol{k}_2')_0\big]+\ldots.+v^M[\ldots]\,.$$

In writing (4.4) we have, of course, suppressed internal coordinate indices. From (3.28) we have immediately [14]

$$(4.5)\qquad \mathscr{D}^0(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2')_0=\frac{f(\varepsilon^0_{k_2})-f(\varepsilon^0_{k_1})}{\omega+\varepsilon^0_{k_2}-\varepsilon^0_{k_1}+i\varepsilon}\,\delta(\boldsymbol{k}_1\boldsymbol{k}_2')\,\delta(\boldsymbol{k}_2\boldsymbol{k}_1')\,,$$

where the functions $f(\omega)$ are defined in (3.3). For $m\neq 0$ the functions $\mathscr{D}^m(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2')_0$ are somewhat more difficult to determine. It is instructive to consider the case for which the w_m of (3.6) are chosen such that

$$(4.6)\qquad w_1=v;\qquad w_m=0\,,\qquad m\geqslant 2\,.$$

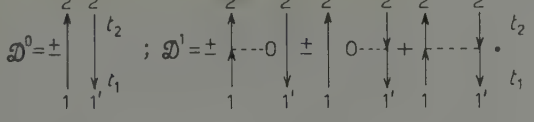
This choice is appropriate to the case when v represents a weak perturbative interaction, since the expansion (3.7) becomes simply an expansion in powers of the interaction. Using (4.6) with (3.9) and proceeding with a tedious but nonetheless straightforward computation, we may determine the various $\mathscr{D}^m$. For $m=1$ we find

$$(4.7)\qquad \mathscr{D}^1(\omega+i\varepsilon;\,\boldsymbol{k}_1\boldsymbol{k}_1',\,\boldsymbol{k}_2\boldsymbol{k}_2')_0=\delta(\boldsymbol{k}_1\boldsymbol{k}_2')\,\delta(\boldsymbol{k}_2\boldsymbol{k}_1')\cdot$$
$$\cdot\sum_{k_2}f(\varepsilon^0_{k_2})\left\{\big[v(\boldsymbol{k}_1\boldsymbol{k}_2;\,\boldsymbol{k}_1\boldsymbol{k}_2)\pm v(\boldsymbol{k}_1\boldsymbol{k}_2;\,\boldsymbol{k}_2\boldsymbol{k}_1)\big]\frac{\mathrm{d}}{\mathrm{d}\varepsilon^0_{k_1}}+\right.$$
$$\left.+\big[v(\boldsymbol{k}_2\boldsymbol{k}_2;\,\boldsymbol{k}_2\boldsymbol{k}_2)\pm v(\boldsymbol{k}_2\boldsymbol{k}_2;\,\boldsymbol{k}_2\boldsymbol{k}_2)\big]\frac{\mathrm{d}}{\mathrm{d}\varepsilon^0_{k_2}}\right\}\cdot\left[\frac{f(\varepsilon^0_{k_2})-f(\varepsilon^0_{k_1})}{\omega+\varepsilon^0_{k_2}-\varepsilon^0_{k_1}+i\varepsilon}\right]+$$
$$+\left[\frac{f(\varepsilon^0_{k_1})-f(\varepsilon^0_{k_1})}{\omega+\varepsilon^0_{k_1}-\varepsilon^0_{k_1}+i\varepsilon}\right]\big[v(\boldsymbol{k}_1\boldsymbol{k}_2;\,\boldsymbol{k}_1'\boldsymbol{k}_2')\pm v(\boldsymbol{k}_1\boldsymbol{k}_2;\,\boldsymbol{k}_2'\boldsymbol{k}_1')\big]\cdot\left[\frac{f(\varepsilon^0_{k_2})-f(\varepsilon^0_{k_2})}{\omega+\varepsilon^0_{k_2}-\varepsilon^0_{k_2}+i\varepsilon}\right].$$

To simplify the subsequent presentation, it is convenient to introduce a diagramatical representation. We shall represent the propagator $G_1^0(11')$ by a solid arrow having its head at the point 1′, tail at 1; we shall represent the symmetrized interaction $v\pm v_{ex}$ by a dashed line. The function $f(\varepsilon^0_k)$, which

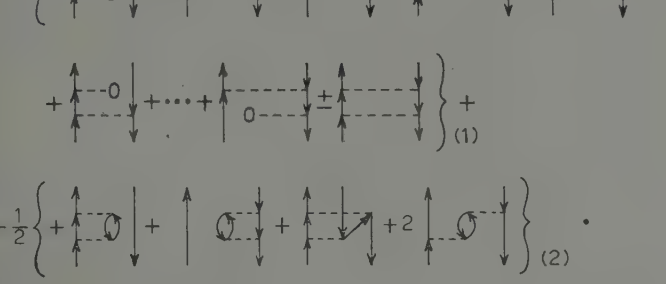
equals $\pm iG_1^0(\boldsymbol{k}_1; 0^-)$, will be represented by a circle or « bubble ». Rewriting (4.5) and (4.7) according to such a scheme, we would have

(4.8)



Using the same prescription, we have the following representation of $\mathscr{D}^2$:

(4.9)



We have separated this expression into two groups of terms since in a certain sense, more intuitive than strictly mathematical, the first group corresponds to an iteration of the first-order terms of $\mathscr{D}^1$ while the second group represents intrinsically second-order effects.

Using $\mathscr{D}^1$ in the $M = 1$ form of equations (4.4), we have after simple algebraic operations

$$(4.10a)\qquad \left\{\omega + i\varepsilon + \varepsilon^1_{k_1'}(k_1 k_1') - \varepsilon^1_{k_1}(k_1 k_1')\right\}\mathscr{D}_1(\omega + i\varepsilon;\, \boldsymbol{k}_1\boldsymbol{k}_1',\, \boldsymbol{k}_2\boldsymbol{k}_2')_0 =$$

$$= \delta(\boldsymbol{k}_1\boldsymbol{k}_2')\,\delta(\boldsymbol{k}_2\boldsymbol{k}_1')\,R_1(k_1k_1')[f(\varepsilon^0_{k_1'}) - f(\varepsilon^0_{k_1})] +$$

$$+ R_1(k_1k_1')[f(\varepsilon^0_{k_1'}) - f(\varepsilon^0_{k_1})]\cdot\sum_{\boldsymbol{k}_3}\sum_{\boldsymbol{k}_3'}[v(\boldsymbol{k}_1\boldsymbol{k}_3;\, \boldsymbol{k}_1\boldsymbol{k}_3') + v]\mathscr{D}_1(\omega + i\varepsilon;\, \boldsymbol{k}_3'\boldsymbol{k}_3,\, \boldsymbol{k}_2\boldsymbol{k}_2')_0,$$

where

$$(4.10b)\qquad R_1(k_1k_1') = \left\{1 - \sum_{\boldsymbol{k}_3} f(\varepsilon^0_{k_3})\left\{[v \pm v]\frac{\mathrm{d}}{\mathrm{d}\varepsilon^0_{k_1'}} + [v \pm v]\frac{\mathrm{d}}{\mathrm{d}\varepsilon^0_{k_1}}\right\}\ln\left[f(\varepsilon^0_{k_1'}) - f(\varepsilon^0_{k_1})\right]\right\}^{-1},$$

$$(4.10c)\qquad \varepsilon_{k_2}(k_1k_1') = \varepsilon^0_{k_2} + R_1(k_1k_1')\cdot\sum_{\boldsymbol{k}_3} f(\varepsilon^0_{k_3})[v(\boldsymbol{k}_2\boldsymbol{k}_3;\, \boldsymbol{k}_2\boldsymbol{k}_3) \pm v]\,.$$

In the event that exchange interactions are negligible and that the direct

interaction is a function only of $(\boldsymbol{r}_1 - \boldsymbol{r}_2)$

$$(4.11) \qquad v_d(\boldsymbol{k}_1\boldsymbol{k}_2;\ \boldsymbol{k}_1'\boldsymbol{k}_2') = \delta(\boldsymbol{k}_1 + \boldsymbol{k}_2;\ \boldsymbol{k}_1' + \boldsymbol{k}_2')\cdot v_d(\tfrac{1}{2}(\boldsymbol{k}_1 - \boldsymbol{k}_2) - \tfrac{1}{2}(\boldsymbol{k}_1' - \boldsymbol{k}_2')),$$

we may solve (4.10) exactly to obtain thereby [15]

$$(4.12a) \qquad \mathscr{D}_1(\omega + i\varepsilon;\ \boldsymbol{k})_0 \equiv \sum_{\boldsymbol{k}_1}\sum_{\boldsymbol{k}_2} \mathscr{D}_1(\omega + i\varepsilon;\ \boldsymbol{k}_1\boldsymbol{k}_2;\ \boldsymbol{k})_0 = \frac{X_1(\omega + i\varepsilon;\ \boldsymbol{k})_0}{1 - v_d(\boldsymbol{k})\cdot X_1(\omega + i\varepsilon;\ \boldsymbol{k})_0},$$

where

$$(4.12b) \qquad X_1(\omega + i\varepsilon;\ \boldsymbol{k})_0 \equiv \sum_{\boldsymbol{k}_1} \left\{ \frac{[f(\varepsilon^0_{\boldsymbol{k}_{1-}}) - f(\varepsilon^0_{\boldsymbol{k}_{1+}})]}{\{1 + \beta v_d(0)[1 \pm f(\varepsilon^0_{\boldsymbol{k}_{1-}}) \pm f(\varepsilon^0_{\boldsymbol{k}_{1+}})]\cdot \sum f(\varepsilon^0_{\boldsymbol{k}_3})\}} \cdot \right.$$

$$\left. \cdot (\omega + i\varepsilon + \varepsilon^0_{\boldsymbol{k}_{1-}} - \varepsilon^0_{\boldsymbol{k}_{1+}})^{-1} \right\} .$$

$$(4.12c) \qquad \xrightarrow[\boldsymbol{k}\to 0]{} \sum_{\boldsymbol{k}_1} \frac{\beta(\boldsymbol{k}\cdot\nabla\varepsilon^0_{\boldsymbol{k}_1})\cdot f(\varepsilon^0_{\boldsymbol{k}_1})[1 \pm f(\varepsilon^0_{\boldsymbol{k}_1})]}{\{1 + \beta v_d(0)\,[1 \pm 2f(\varepsilon^0_{\boldsymbol{k}_1})]\cdot \sum f(\varepsilon^0_{\boldsymbol{3}})\}} (\omega + i\varepsilon) - \boldsymbol{k}\cdot\nabla\varepsilon^1_{\boldsymbol{k}_1})^{-1}.$$

Expression (4.12) or a simple modification thereof [16] has been studied for a number of interesting cases. It has been found to be particularly relevant in treating collective density modes of low-temperature degenerate *F-D* and *B-E* systems [17].

With respect to medium temperature *M-B* systems — systems of molecules forming solids, liquids, and gases at ordinary temperatures and pressures, for example — one finds that, under the strong-force high-density conditions which obtain in solids and liquids, it is conceivable that a higher-order form of an equation of the type (4.12) will predict a collective sound mode for such systems. However, it is noteworthy that sound propagation in weakly interacting dilute *M-B* gases is *not* described by a $\mathscr{D}_1$-equation. To better understand this, let us consider (4.12) in the long wavelength limit

---

[15] A similar result obtains of course in the event that the exchange interaction, though possibly not negligible, may be approximated for the relevant range of momenta by a constant.

[16] For hard cores, for example, $v(k)$ must be replaced by some sort of scattering matrix. For the electron gas the singular $v(0)$ must be compensated by a neutralizing charge distribution whose effects would appear in the formalism as an external field.

[17] The phenomenon of « zero sound » in *F-D* systems has been treated by a number of authors; a recent treatment containing references to older literature is the work of K. Gottfried and L. Pičman: *Phys. Rev.* (to be published). A related phenomenon, electron plasma oscillation, has also been extensively treated; we shall cite simply the article by K. Sawada, K. A. Brueckner, N. Fukuda and R. Brout: *Phys. Rev.*, **108**, 507 (1957). Although the analysis is complicated by fluctuation considerations, one may also show that the $\mathscr{D}_1$-equation is relevant in treating the phonon spectrum of a condensed *B-E* gas.

($\boldsymbol{k} \to 0$), retaining only terms of highest order in the fugacity $\exp[-\alpha]$ and choosing

$$v_a(k) = v_0 \frac{k_0^2}{k_0^2 + k^2} \xrightarrow{k \to 0} v_0 > 0; \qquad \varepsilon^0_{k_1} = \frac{k_1^2}{2m}. \tag{4.13}$$

We have for this case

$$\mathscr{D}_1(\omega + i\varepsilon; \boldsymbol{k})_0 \to \frac{\left\{\beta n_0 \int_{-1}^{1} d\mu \int_0^\infty d\chi \, \frac{\mu^2 \chi^4 \exp[-\chi^2]}{(\bar{\omega} + i\varepsilon)^2 - \mu^2 \chi^2}\right\}}{\left\{2(1 + \beta v_0 n_0) \int_0^\infty d\chi \, \chi^2 \exp[-\chi^2] - \beta n_0 v_0 \int_{-1}^{1} d\mu \int_0^\infty d\chi \, \frac{\mu^2 \chi^4 \exp[-\chi^2]}{(\bar{\omega} + i\varepsilon)^2 - \mu^2 \chi^2}\right\}}, \tag{4.14a}$$

where we have introduced $\bar{\omega}$ such that

$$\omega^2 = \frac{2k^2}{\beta m} \bar{\omega}^2. \tag{4.14b}$$

The sound resonance, if it exists, will be associated with a zero of the real part of the denominator (ω real) in (4.14). One may readily determine that for $\bar{\omega}$ real the maximum value of the ratio

$$\left\{\int_{-1}^{1} d\mu \int_0^\infty d\chi \, \chi^4 \mu^2 \exp[-\chi^2] \frac{P}{\bar{\omega}^2 - \mu^2 \chi^2}\right\} \cdot \left\{2 \int_0^\infty d\chi \, \chi^2 \exp[-\chi^2]\right\}^{-1}, \tag{4.15}$$

is on the order of unity. If follows from this fact that a sound resonance could occur in (4.14) only if

$$\beta v_0 n \gtrsim 1 \quad \text{or} \quad \frac{\lambda^3 n}{V} \cdot \frac{a}{\lambda} \gtrsim 1, \tag{4.16}$$

where in the second expression we have introduced a scattering length a and the thermal wavelength $\lambda = \sqrt{4\pi\beta}$. The condition (4.16) is strongly violated by a weakly-interacting dilute M-B gas. We shall see later that, provided the sound wavelength is sufficiently long ($\boldsymbol{k} \to 0$), the interaction strength $v_0 > 0$ may be made arbitrarily small and the sound propagation in the gas will still be maintained. This, of course, would not be admitted by (4.16). With respect to liquidis and solids for which (4.16) is easily met, we note that is is not reasonable to restrict ourselves only to the first order approximation

(4.12). Moreover, it is well known that the simple random-phase approximation (4.12) exhibits inadmissible instabilities if one attempts its solution with the attractive binding interaction of solids ad liquids.

In actual fact the sound resonance associated with weakly interacting dilute *M-B* gases, « perfect » gases, is a second-order phenomenon associated with the occurrence of damping in the single-particle Green's function G_1. We shall devote the remainder of this work to the consideration of the second order case. In treating that case it is convenient to utilize the separation of terms indicated in (4.9), writing (4.4) in the form

$$\mathscr{D}_2 = \mathscr{D}^0 + [\nu\mathscr{D}^1 + \nu^2\mathscr{D}^2_{(1)} - \nu^2\mathscr{D}^1(\mathscr{D}^0)^{-1}\mathscr{D}^1](\mathscr{D}^0)^{-1}\mathscr{D}_2 + \nu^2\mathscr{D}^2_{(2)}(\mathscr{D}^0)^{-1}\mathscr{D}_2 . \tag{4.17}$$

Substituting explicit expressions for $\mathscr{D}^0$, $\mathscr{D}^1$, and $\mathscr{D}^2_{(1)}$ into (4.17), we obtain after simple operations

$$\begin{aligned}
&\{\omega + i\varepsilon + \varepsilon^2_{k_1'}(k_1k_1') - \varepsilon^2_{k_1}(k_1k_1')\}\,\mathscr{D}_2(\omega + i\varepsilon;\, \boldsymbol{k}_1\boldsymbol{k}_1',\, \boldsymbol{k}_2\boldsymbol{k}_2')_0 \\
&= \delta(\boldsymbol{k}_1\boldsymbol{k}_2')\,\delta(\boldsymbol{k}_2\boldsymbol{k}_1')\cdot R_2(k_1k_1')[f(\varepsilon^0_{k_1'}) - f(\varepsilon^0_{k_1})] + \\
&+ R_2(k_1k_1')\cdot[f(\varepsilon^0_{k_1'}) - f(\varepsilon^0_{k_1})]\cdot\sum_{\boldsymbol{k}_3}\sum_{\boldsymbol{k}_3'}[v(\boldsymbol{k}_1\boldsymbol{k}_3;\, \boldsymbol{k}_1'\boldsymbol{k}_3') \pm v]\cdot\mathscr{D}(\omega + i\varepsilon;\, \boldsymbol{k}_3'\boldsymbol{k}_3,\, \boldsymbol{k}_2\boldsymbol{k}_2')_0 + \\
&+ R_2(k_1k_1')\cdot[\omega + i\varepsilon + \varepsilon^0_{k_1'} - \varepsilon^0_{k_1}]\cdot\mathscr{D}^2_{(2)}(\mathscr{D}^0)^{-1}\mathscr{D}_2(\omega + i\varepsilon;\, \boldsymbol{k}_1\boldsymbol{k}_1',\, \boldsymbol{k}_2\boldsymbol{k}_2')_0 ,
\end{aligned} \tag{4.18a}$$

where

$$\begin{aligned}
R_2(k_1k_1') = \Bigg[1 - \sum_3\Bigg\{1 + \sum_4 f(\varepsilon^0_4)[v \pm v]\frac{d}{d\varepsilon^0_3}\Bigg\} f(\varepsilon^0_3)\cdot\Bigg\{[v \pm v]\frac{d}{d\varepsilon^0_{1'}} + \\
+ [v \pm v]\frac{d}{d\varepsilon^0_1}\Bigg\}\ln[f(\varepsilon^0_{1'}) - f(\varepsilon^0_1)] - \tfrac{1}{2}[f(\varepsilon^0_{1'}) - f(\varepsilon^0_1)]^{-1}\cdot \\
\cdot\Bigg\{\sum_3 f(\varepsilon^0_3)\Bigg\{[v \pm v]\frac{d}{d\varepsilon^0_{1'}} + [v \pm v]\frac{d}{d\varepsilon^0_1}\Bigg\}\Bigg\}^2[f(\varepsilon^0_{1'}) - f(\varepsilon^0_1)] + \\
+ \Bigg\{\sum_3 f(\varepsilon^0_3)\Bigg\{[v \pm v]\frac{d}{d\varepsilon^0_{1'}} + [v \pm v]\frac{d}{d\varepsilon^0_1}\Bigg\}\ln[f(\varepsilon^0_{1'}) - f(\varepsilon^0_1)]\Bigg\}^2\Bigg]^{-1},
\end{aligned} \tag{4.18b}$$

$$\begin{aligned}
\varepsilon^2_{k_2}(k_1k_1') = \varepsilon^0_{k_2} + R_2(k_1k_1')\cdot\sum_3\Bigg[1 + \sum_4 f(\varepsilon^0_4)\Bigg\{[v \pm v]\frac{d}{d\varepsilon^0_3} + \\
+ \Bigg\{[v \pm v]\frac{d}{d\varepsilon^0_{1'}} + [v \pm v]\frac{d}{d\varepsilon^0_1}\Bigg\}\ln[f(\varepsilon^0_{1'}) - f(\varepsilon^0_1)]\Bigg\}\Bigg]\cdot f(\varepsilon^0_3)[v(\boldsymbol{k}_2\boldsymbol{k}_3;\, \boldsymbol{k}_2\boldsymbol{k}_3) \pm v] .
\end{aligned} \tag{4.18c}$$

Under the conditions for which (4.12) obtains, the above result may be simplified somewhat. Rather than pursue the analogy with the $\mathscr{D}_1$-case further,

however, we shall shift our attention to the component $\mathscr{D}^2_{(2)}$ which has only been symbolically represented in (4.18).

A detailed treatment of $\mathscr{D}^2_{(2)}$, although straightforward, would require far too much space since the various expressions are not readily written in a concise manner. For this reason and in order not to obscure the essential features of the calculation within a maze of algebra, we shall make a number of simplifications which are particularly appropriate to the treatment of the long wavelength sound modes in a perfect gas. To this end we note that with the choice (4.6) the integral equation (3.20) yields the following second-order expression for the single-particle Green's function G_1:

$$(4.19a)\qquad G_1(\omega+i\varepsilon;\,\boldsymbol{k}_1)_0^{-1}=\omega+i\varepsilon-\varepsilon^0_{k_1}-\sum_3\left\{1+\sum_4 f(\varepsilon^0_4)[v\pm v]\frac{\mathrm{d}}{\mathrm{d}\varepsilon^0_3}\right\}\cdot$$

$$\cdot f(\varepsilon^0_3)[v(\boldsymbol{k}_1\boldsymbol{k}_3;\,\boldsymbol{k}_1\boldsymbol{k}_3)\pm v]-\tfrac{1}{2}\sum_2\sum_3\sum_4\left\{f(\varepsilon^0_2)[1\pm f(\varepsilon^0_3)\pm f(\varepsilon^0_4)]\mp f(\varepsilon^0_3)\,f(\varepsilon^0_4)\right\}\cdot$$

$$\cdot\frac{[v(\boldsymbol{k}_1\boldsymbol{k}_2;\,\boldsymbol{k}_3\boldsymbol{k}_4)\pm v]\,[v(\boldsymbol{k}_3\boldsymbol{k}_4;\,\boldsymbol{k}_1\boldsymbol{k}_2)\pm v]}{(\omega+i\varepsilon+\varepsilon^0_2-\varepsilon^0_3-\varepsilon^0_4)},$$

$$(4.19b)\qquad \underset{(\omega\ \mathrm{real})}{\approx}\omega+i\varepsilon-\varepsilon^0_{k_1}-\sum_3\left\{1+\sum_4 f(\varepsilon^0_4)v\pm v\right]\frac{\mathrm{d}}{\mathrm{d}\varepsilon^0_3}\Big\}f(\varepsilon^0_3)[v\pm v]-$$

$$-\tfrac{1}{2}\sum_2\sum_3\sum_4\left\{f(\varepsilon^0_2)[1\pm f(\varepsilon^0_3)\pm f(\varepsilon^0_4)]\mp f(\varepsilon^0_3)f(\varepsilon^0_4)\right\}\cdot$$

$$\cdot[v\pm v]\,[v\pm v]\left\{\frac{P}{\varepsilon^0_1+\varepsilon^0_2-\varepsilon^0_3-\varepsilon^0_4}-i\pi\delta(\varepsilon^0_1+\varepsilon^0_2-\varepsilon^0_3-\varepsilon^0_4)\right\}.$$

The approximation (4.19*b*) for ω real is excellent under perfect gas conditions but is less valid in those cases for which the $\mathscr{D}_1$-equation (4.12) yields a collective resonance. The second-order term of (4.19*a*) appears in conjunction with propagators G_1^0 in the first two terms (cfr. (4.9)) of $\mathscr{D}^2_{(2)}$. All of the terms of $\mathscr{D}^2_{(2)}$ are closely related in structure; in each of those terms we shall make the approximation equivalent to that indicated in (4.19). Since the last two terms of $\mathscr{D}^2_{(2)}$ are not diagonal in k-space, there exists a slight ambiguity in the frequency approximation to be employed. We shall resolve that ambiguity by requiring that the approximation maintain the structural similarity of the $\mathscr{D}^2_{(2)}$ terms. (This is tantamount to asserting that the relevant frequency is that of the «driving» function $(\mathscr{D}^0)^{-1}\mathscr{D}^2$ on the right hand side of (4.17)). In fact, this choice may be shown to be the only one compatible with the time-invariance conditions previously discussed. As in (4.19) the postulated approximation will generate in each term a real and an imaginary component. The real components could be absorbed into a $\mathscr{D}_1$-type equation, as were the components of $\mathscr{D}^2_{(1)}$ in (4.18). However, since the expres-

sions involved are extremely unwieldy, we shall further simplify our analysis by omitting all real components of $(\mathscr{D}^0)^{-1}[\mathscr{D}^2 - \mathscr{D}^1(\mathscr{D}^0)^{-1}\mathscr{D}^1](\mathscr{D}^0)^{-1}$ in (4.17). This simplification will not destroy the time-invariance properties since the invariance relations must clearly hold for the real and imaginary components separately. Thus we shall maintain order-v accuracy in the real components and order-v^2 accuracy in the imaginary components. Doing that, we are led to consider the equation

$$(4.20)\quad \{\omega + i\varepsilon + \varepsilon^1_{k_{1-}}(\boldsymbol{k}_{1+}, \boldsymbol{k}_{1-}) - \varepsilon^1_{k_{1+}}(k_{1+}, k_{1-})\}\mathscr{D}(\omega + i\varepsilon; \boldsymbol{k}_1\boldsymbol{k}_2; \boldsymbol{k})_0 = R_1(\boldsymbol{k}_{1+}, \boldsymbol{k}_{1-})\cdot$$
$$\cdot\{[f(\varepsilon^0_{k_{1-}}) - f(\varepsilon^0_{k_{1+}})]\cdot\{\delta(\boldsymbol{k}_1\boldsymbol{k}_2) + \sum_3[v(\boldsymbol{k}_{1+}, \boldsymbol{k}_{3-}; \boldsymbol{k}_{1-}, \boldsymbol{k}_{3+}) \pm v]\mathscr{D}(\omega + i\varepsilon; \boldsymbol{k}_3\boldsymbol{k}_2; \boldsymbol{k})_0\} -$$
$$- i\sum_3[K_{12}(\boldsymbol{k}_3\boldsymbol{k}_1)_{\boldsymbol{k}}\mathscr{D}(\omega + i\varepsilon; \boldsymbol{k}_1\boldsymbol{k}_2; \boldsymbol{k})_0 - K_{34}(\boldsymbol{k}_1\boldsymbol{k}_3)_{\boldsymbol{k}}\mathscr{D}(\omega + i\varepsilon; \boldsymbol{k}_3\boldsymbol{k}_2; \boldsymbol{k})_0]\}$$

valid for ω real . The functions ε^1, R_1 are defined in (4.10). The functions $K(\boldsymbol{k}_1\boldsymbol{k}_2)_{\boldsymbol{k}}$ are the following real even functions of $\boldsymbol{k}$:

$$(4.21a)\quad K_{12}(\boldsymbol{k}_1\boldsymbol{k}_2)_{\boldsymbol{k}} = \frac{\pi}{2}\sum_3\sum_4[1 \pm f(\varepsilon^0_{k_1})]f(\varepsilon^0_{k_3})f(\varepsilon^0_{k_4})\{[f(\varepsilon^0_{k_{2+}})]^{-1}\cdot[v(\boldsymbol{k}_1\boldsymbol{k}_{2+}; \boldsymbol{k}_3\boldsymbol{k}_4) \pm v]\cdot$$
$$\cdot[v(\boldsymbol{k}_3\boldsymbol{k}_4; \boldsymbol{k}_1\boldsymbol{k}_{2+}) \pm v]\cdot\delta(\varepsilon^0_{k_1} + \varepsilon^0_{k_{2+}} - \varepsilon^0_{k_3} - \varepsilon^0_{k_4}) + [f(\varepsilon^0_{k_{2-}})]^{-1}\cdot$$
$$\cdot[v(\boldsymbol{k}_1\boldsymbol{k}_{2-}; \boldsymbol{k}_3\boldsymbol{k}_4) \pm v][v(\boldsymbol{k}_3\boldsymbol{k}_4; \boldsymbol{k}_1\boldsymbol{k}_{2-}) \pm v]\cdot\delta(\varepsilon^0_{k_1} + \varepsilon^0_{k_{2-}} - \varepsilon^0_{k_3} - \varepsilon^0_{k_4})\};$$

$$(4.21b)\quad K_{34}(\boldsymbol{k}_i\boldsymbol{k}_2)_{\boldsymbol{k}} = \frac{\pi}{2}\sum_3\sum_4 f(\varepsilon^0_{k_3})f(\varepsilon^0_{k_4})\Big\{\exp[\alpha + \beta\varepsilon^0_{k_3}]\cdot$$
$$\cdot[v(\boldsymbol{k}_3\boldsymbol{k}_{2+}; \boldsymbol{k}_{1+}\boldsymbol{k}_4) \pm v][v(\boldsymbol{k}_{1-}\boldsymbol{k}_4; \boldsymbol{k}_3\boldsymbol{k}_{2-}) \pm v]\,2\begin{Bmatrix} f(\varepsilon^0_{k_{1-}}) \\ f(\varepsilon^0_{k_{2-}}) \end{Bmatrix}\delta(\varepsilon^0_{k_{1-}} + \varepsilon^0_{k_4} - \varepsilon^0_{k_3} - \varepsilon^0_{k_{2-}}) +$$
$$+ \frac{f(\varepsilon^0_{k_{1+}})}{f(\varepsilon^0_{k_{2+}})}\delta(\varepsilon^0_{k_{1+}} + \varepsilon^0_{k_4} - \varepsilon^0_{k_3} - \varepsilon^0_{k_{2+}})\Big\} - [v(\boldsymbol{k}_{1+}\boldsymbol{k}_{2-}; \boldsymbol{k}_3\boldsymbol{k}_4) \pm v][v(\boldsymbol{k}_3\boldsymbol{k}_4; \boldsymbol{k}_{2+}\boldsymbol{k}_{1-}) \pm v]\cdot$$
$$\cdot\left\{\frac{[1 \pm f(\varepsilon^0_{k_{1-}})]}{f(\varepsilon^0_{k_{2+}})}\delta(\varepsilon^0_{k_{2+}} + \varepsilon^0_{k_{1-}} - \varepsilon^0_{k_3} - \varepsilon^0_{k_4}) + \frac{[1 \pm f(\varepsilon^0_{k_{1+}})]}{f(\varepsilon)^0_{k_{2-}}}\delta(\varepsilon^0_{k_{2-}} + \varepsilon^0_{k_{1+}} - \varepsilon^0_{k_3} - \varepsilon^0_{k_4})\right\}\Big\}.$$

In contrast to the real components previously neglected the functions (4.21) remain finite in the limit $\boldsymbol{k} \to 0$. If we define

$$(4.22)\qquad \Gamma(\boldsymbol{k}_1)_{\boldsymbol{k}} \equiv \sum_{k_3} K_{12}(\boldsymbol{k}_3\boldsymbol{k}_1)_{\boldsymbol{k}},$$

then in the limit $\boldsymbol{k} \to 0$ the function $\Gamma(\boldsymbol{k}_1)_{\boldsymbol{k}}$ is to the lowest order in the fugacity equal to the half-width to be associated with the approximate single-

particle function (4.19*b*) in the representation (1.13). The $\boldsymbol{k}=0$ functions have the following two important additional properties:

$$\Gamma(\boldsymbol{k}_1)_0 = \sum_{\boldsymbol{k}_3} K_{12}(\boldsymbol{k}_3\,\boldsymbol{k}_1)_0 = \sum_{\boldsymbol{k}_3} K_{34}(\boldsymbol{k}_3\,\boldsymbol{k}_1)_0 ; \tag{4.23a}$$

$$\begin{aligned} Y_{\mathscr{A}} &= \sum_{\boldsymbol{k}_1} \mathscr{A}(\boldsymbol{k}_1)\cdot \sum_{\boldsymbol{k}_3} \left[K_{12}(\boldsymbol{k}_3\boldsymbol{k}_1)_0\,\mathscr{D}(\boldsymbol{k}_1) - K_{34}(\boldsymbol{k}_1\boldsymbol{k}_3)_0\,\mathscr{D}(\boldsymbol{k}_3)\right] = \\ &= \sum_{\boldsymbol{k}_1}\sum_{\boldsymbol{k}_2}\sum_{\boldsymbol{k}_3}\sum_{\boldsymbol{k}_4} \left[\mathscr{A}(\boldsymbol{k}_1) + \mathscr{A}(\boldsymbol{k}_2) - \mathscr{A}(\boldsymbol{k}_3) - \mathscr{A}(\boldsymbol{k}_4)\right]\cdot\pi\cdot \frac{\left[1 \pm f(\varepsilon^0_{k_1})\right] f(\varepsilon^0_{k_3})\, f(\varepsilon^0_{k_4})}{f(\varepsilon^0_{k_2})}\cdot \\ &\quad \cdot\left[v(\boldsymbol{k}_1\boldsymbol{k}_2;\,\boldsymbol{k}_3\boldsymbol{k}_4) \pm v\right]\left[v(\boldsymbol{k}_3\boldsymbol{k}_4;\,\boldsymbol{k}_1\boldsymbol{k}_2) \pm v\right]\cdot\delta(\varepsilon^0_{k_1} + \varepsilon^0_{k_2} - \varepsilon^0_{k_3} - \varepsilon^0_{k_4})\cdot\mathscr{D}(\boldsymbol{k}_2)\,. \end{aligned} \tag{4.23b}$$

The function $Y_{\mathscr{A}}$ is zero, independent of $\mathscr{D}(\boldsymbol{k}_2)$, for the following important choices of $\mathscr{A}(\boldsymbol{k}_1)$:

$$\left\{\begin{array}{lll} (a) & \mathscr{A}(\boldsymbol{k}_1) = \text{constant} & : [N, H] = 0; \\ (b) & \mathscr{A}(\boldsymbol{k}_1) = \boldsymbol{k}_1 & : [\boldsymbol{P}, H] = 0; \\ (c) & \mathscr{A}(\boldsymbol{k}_1) = \varepsilon & : [H, H] = 0\,. \end{array}\right. \tag{4.24}$$

If the interaction v is diagonal in other coordinates as well as the total momentum, there will of course be additional invariants. The various choices correspond in the present order of calculation to the « summational invariants » to be associated, as was discussed in Section **3**, with the indicated zero commutators.

It is apparent from the preceding remarks that (4.20) is a type of Boltzmann equation. By continuing to higher-order forms of (4.4), using (4.6) or some more sagacious choice of the w_m, one could in principle extend that equation to arbitrarily high accuracy, maintaining the necessary invariants at each stage. The form (4.20) is sufficient to describe sound propagation in a perfect gas, and we shall consider that specific case in the next section. In the process we shall develop approximations which are appropriate for studying more general features of the solutions of (4.20) than merely the sound modes.

5. – Sound propagation in a nearly-perfect gas.

The usual methods for obtaining the sound modes from a Boltzmann equation do little to clarify the microscopic propagation mechanism nor do they indicate in any precise manner the limits of validity of the procedures. In treating the long wavelength ($\boldsymbol{k} \to 0$) sound modes of a nearly-perfect gas, we propose to circumvent these two objections by treating quantitatively

the K-function kernels of the integral equation (4.20). Unfortunately, even for relatively simple choices of v and ε_k^0, the functions (4.21) are much too complicated to admit the normal methods of analytic solution. In order to pursue this approach further, we must therefore introduce approximations to the K-functions.

To simplify our considerations, let us assume that the kernels (4.21) may be replaced by their $\boldsymbol{k} = 0$ forms. This is by no means an essential step, but it does greatly simplify the mathematical forms. The order-k^2 terms neglected in the $\boldsymbol{k} = 0$ approximation would first enter in the terms describing the damping of the sound resonance. With the $\boldsymbol{k} = 0$ assumption we may by virtue of (4.23a) replace K_{12} by K_{34}. Further, we may write K_{34} in the form

$$K_{34}(\boldsymbol{k}_1 \boldsymbol{k}_2)_0 = \sum_l P_l\left(\cos \frac{\boldsymbol{k}_1}{\boldsymbol{k}_2}\right) \cdot K_l(k_1 k_2), \tag{5.1}$$

where the $K_l(k_1, k_2)$ are functions only of the magnitudes k_1, k_2. Using (4.21) we may easily verify the symmetry property

$$K_l(k_1 k_2) \cdot [f(\varepsilon_{k_2}^0)]^2 \exp [\alpha + \beta \varepsilon_{k_2}^0] = K_l(k_2 k_1) \cdot [f(\varepsilon_{k_1}^0)]^2 \exp [\alpha + \beta \varepsilon_{k_1}^0],$$

from which we are led to approximate K_l by the partial sum

$$K_l(k_1 k_2) = [f(\varepsilon_{k_2}^0)]^{-2} \exp [-\alpha - \beta \varepsilon_{k_2}^0] \cdot \sum_{n=1}^{N_l} g_{ln}(k_1) g_{ln}(k_2) . \tag{5.2}$$

The functions $g_{ln}(k)$ must be chosen so as to satisfy the magnitude requirement (4.23a)

$$\Gamma(k_1) = \sum_{\boldsymbol{k}_3} K_{34}(\boldsymbol{k}_3 \boldsymbol{k}_1)_0 = [f(\varepsilon_{k_1}^0)]^{-2} \exp [-\alpha - \beta \varepsilon_{k_1}^0] \cdot \sum_n g_{0n}(k_1) \cdot [\sum_{\boldsymbol{k}_3} g_{0n}(k_3)]; \tag{5.3a}$$

and the invariance conditions implied by (4.23b) and (4.24):

$$[\boldsymbol{P}, H] = 0 : k_1 \sum_n g_{0n}(k_1) \cdot [\sum_{\boldsymbol{k}_3} g_{0n}(k_3)] = \tfrac{1}{3} \sum_n g_{1n}(k_1) \cdot [\sum_{\boldsymbol{k}_3} k_3 g_{1n}(k_3)]; \tag{5.3b}$$

$$[H, H] = 0 : \varepsilon_{k_1}^0 \sum_n g_{0n}(k_1) \cdot [\sum_{\boldsymbol{k}_3} g_{0n}(k_3)] = \sum_n g_{0n}(k_1) \cdot [\sum_{\boldsymbol{k}_3} \varepsilon_{k_3}^0 g_{0n}(k_3)] . \tag{5.3c}$$

The $[N, H] = 0$ condition is trivial since we have already used (4.23a). We remark that, except in the classical limit where angular momentum behaves like an ordinary vector, the condition corresponding to $[\boldsymbol{J}, H] = 0$ might be very difficult to implement. If desired, one could compute $\boldsymbol{k} \neq 0$ modifications to (5.1) and (5.2); we shall not do that.

In order to simplify the K-independent terms of (4.20), let us assume as in (4.12) that exchange interactions are negligible and that the direct interaction (4.11) obtains. Neglecting the $0(k^3)$ terms of $\varepsilon^1_{k_1=k}$ and the $0(k^2)$ terms of $R_1(k_{1+}, k_{1-}) \to R_1(k_1)$, using the expressions (5.2), and summing (4.20) over k_2, we obtain a simplified integral equation appropriate to the $k \to 0$ limit.

$$(5.4)\quad \{\omega + i\varepsilon - k\cdot\nabla\varepsilon^0_{k_1} + iR_1\Gamma(k_1)\}\cdot\sum_l P_l\left(\cos\begin{matrix}k_1\\ k\end{matrix}\right)\cdot\mathscr{D}_l(\omega+i\varepsilon;\, k_1k)_0 =$$

$$= \{1+\beta v_d(0)[1\pm 2f(\varepsilon^0_{k_1})]\cdot\sum f(\varepsilon^0_{k_3})\}^{-1}\cdot$$

$$\cdot\Bigg\{(k\cdot\nabla\varepsilon^0_{k_1})\beta f(\varepsilon^0_{k_1})\,[1\pm f(\varepsilon^0_{k_1})]\,\{1+v_d(k)\cdot\sum_{k_3}\mathscr{D}_0(\omega+i\varepsilon;\, k_3k)_0\}+$$

$$+ i\sum_l \frac{P_l\left(\cos\begin{matrix}k_1\\ k\end{matrix}\right)}{2l+1}\sum_n g_{ln}(k_1)\cdot\sum_{k_3}\frac{g_{ln}(k_3)}{f(\varepsilon^0_{k_3})[1\pm f(\varepsilon^0_{k_3})]}\mathscr{D}_l(\omega+i\varepsilon;\, k_3k)_0\Bigg\}.$$

In (5.4) we have introduced the representation

$$(5.5)\quad \mathscr{D}(\omega+i\varepsilon;\, k_1k) \equiv \sum_{k_2}\mathscr{D}(\omega+i\varepsilon;\, k_1k_2;\, k) = \sum_l P_l\left(\cos\begin{matrix}k_1\\ k\end{matrix}\right)\mathscr{D}_l(\omega+i\varepsilon;\, k_1k).$$

We are interested in the ω-plane singularities of (4.4) in the region $\omega \sim k|\nabla\varepsilon|_{\text{ave}}$. Since $\Gamma(k_1)_0$ is independet of k, it follows that in the limit $k \to 0$ we may utilize the expansion

$$(5.6)\quad \{\omega+i\varepsilon-k\cdot\nabla\varepsilon^0_{k_1}+iR_1(k_1)\Gamma(k_1)\}^{-1} = \frac{1}{iR_1\Gamma(k_1)}\left\{1+i\frac{\omega+i\varepsilon-k\cdot\nabla\varepsilon_{k_1}}{R_1\Gamma_1(k_1)}+\cdots\right\},$$

in investigating those singularities.

The essential simplification which we shall make to reduce (4.20) or (5.4) to a tractable form is to represent the functions K_l by only a few terms of the sums (5.2), but enough terms such that the conditions (5.3) are met. In principle, this approximation could be extended to arbitrary accuracy since the integral equation (5.4), having a sum of separable functions as its kernel, is soluble. In practice, however, it is reasonable to believe that for most purposes a few well chosen terms will suffice. For simplicity in presentation we shall use the minimum number of terms compatible with (5.3). Thus, we shall take

$$(5.7)\quad K_l(k_1k_2)\cdot[f(\varepsilon^0_{k_2})]^{-2}\exp[\alpha+\beta\varepsilon^0_{k_2}] = \begin{cases} g_{01}(k_1)g_{01}(k_2)+g_{02}(k_1)g_{02}(k_2) & \text{for } l=0;\\ g_1(k_1)g_1(k_1) & \text{for } l=1;\\ 0 & \text{for } l\geqslant 2.\end{cases}$$

We shall indicate later the nature of the modifications induced in the results if one were to supplement the choice (5.7) by several additional terms.

Substituting the expressions (5.7) into (5.4), we are led to a set of four coupled linear equations whose solutions describe the collective modes of the system in the relevant ω-plane region. In matrix form those equations are

$$(5.8)\quad [M]\begin{Bmatrix} \mathscr{D}(\omega+i\varepsilon,\,k)_0 \\ \sum_{k_1} \dfrac{g_{01}\mathscr{D}_0(\omega+i\varepsilon;\,k_1k)_0}{f(1\pm f)} \\ \sum_{k_1} \dfrac{g_{02}\mathscr{D}_0(\omega+i\varepsilon;\,k_1k)_0}{f(1\pm f)} \\ \sum_{k_1} \dfrac{g_1\mathscr{D}_1(\omega+i\varepsilon;\,k_1k)_0}{f(1\pm f)} \end{Bmatrix} = \sum_{k_1} \frac{\beta}{\Gamma^2 R_1^3} \begin{Bmatrix} +k^2|\nabla\varepsilon|^2(1\pm f)+\dots \\ +k^2|\nabla\varepsilon|^2 g_{01}+\dots \\ +k^2|\nabla\varepsilon|^2 g_{02}+\dots \\ +ikg_1|\nabla\varepsilon|3\Gamma R_1\left\{1+i\,\dfrac{\omega}{\Gamma R_1}+\dots\right\} \end{Bmatrix},$$

where the matrix $[M]$ has the specific form indicated in the Appendix. The ω-plane singularities in the sound-resonance region are determined by the zeros of $\det[M]$. One may show by tedious but direct computation that

$$(5.9)\quad \det[M] = -B_1B_2 - i\omega\{B_1B_3+B_2B_4\} + \Big\{\omega^2[B_3B_4+B_1(\dots)+B_2(\dots)] -$$
$$-\frac{k^2}{3^2}\{[B_{1a}B_{2b}^2+B_{1b}B_{6a}^2-2B_{1c}B_{2a}B_{cb}]+\beta v_d(k)C_3[B_{1a}B_{cb}C_2+$$
$$+B_{1b}B_{6a}C_1-B_{1c}(B_{6a}C_2+B_{cb}C_1)]+B_1(\dots)+B_2(\dots)\}\Big\}+$$
$$+i\omega\Big\{\omega^2[B_3B_2+B_{1j}(\dots)+B_2(\dots)]-\frac{k^2}{3^2}\{[B_{4a}B_{cb}^2+B_{4b}B_{ca}^2-2B_{4c}B_{6a}B_{cb}]+$$
$$+\beta v_d(k)C_3[B_{4a}B_{cb}C_2+B_{4b}B_{ca}C_1-B_{4c}(B_{6a}C_2+B_{6b}C_1)]+B_{1j}(\dots)+B_2(\dots)\}\Big\}+\dots,$$

where the ω, k-independent functions B_j, C_j have been tabulated in the Appendix. Recall that the $[N, H]=0$ condition has been built into the structure of (5.4) provided (5.3a) obtains. Using (5.3a), we may easily verify that

$$(5.10)\qquad B_1 = 0;\ [N, H] = 0\,.$$

This condition, although to the lowest order in ω/Γ it generates a pole at $\omega=0$, is not sufficient to produce sound propagation.

If we impose the two conditions (5.3a) and (5.3b), we find

$$(5.11)\qquad B_1 = B_2 = 0;\ [N, H] = [\boldsymbol{P}, H] = 0.$$

Using this result in (5.9), we find after some manipulation that to the lowest

order in ω/Γ there exists a (Newton) sound resonance at $\omega = \pm\omega_0$:

$$(5.12a)\quad \mathscr{D}(\omega + i\varepsilon;\, k)_0 = \frac{\beta k}{3}\cdot\frac{[\sum|\nabla\varepsilon|k_1R_1^{-1}f(1\pm f)][\sum|\nabla\varepsilon|k_1f(1\pm f][\sum f(1\pm f)]}{[\sum k_1^2R_1^{-1}f(1\pm f)][\sum f(1\pm f)R_1^{-1}][(\omega+i\varepsilon)^2-\omega_0^2]};$$

$$(5.12b)\quad \omega_0^2 = \frac{k^2}{3}\cdot$$

$$\cdot\frac{[\sum|\nabla\varepsilon|k_1R_1^{-1}f(1\pm f)\{[\sum|\nabla\varepsilon|R_1^{-1}k_1f(1\pm f)]+\beta v_d(k)[\sum f(1\pm f)][\sum k_1|\nabla\varepsilon|f(1\pm f)]\}}{[\sum k_1^2R_1^{-1}(1\pm f)][\sum f(1\pm f)R_1^{-1}]}.$$

Taking into account the next order term in ω/Γ one finds that the resonance half-width is on the order of $k^2(|\nabla\varepsilon|/\Gamma)_{\text{ave}}$. This is in fact a typical result: the frequency ω_0 depends linearly upon k; the half-width depends upon the square of k and the inverse of the « single-particle » width $\Gamma(k_1)$. Observing that

$$(5.13)\quad \sum_{k_1} k_1|\nabla\varepsilon|f(1\pm f) = \frac{3}{\beta}\sum_{k_1} f(\varepsilon^0_{k_1}),$$

we may reduce (5.12) in the long-wavelength ($k \to 0$) low-density ($\exp[-\alpha] \to 0$; R_1 independent of k_1) limit to

$$(5.14a)\quad \mathscr{D}(\omega+i\varepsilon;\, k)_0 = \frac{k^2\cdot\sum f}{m[1+v_d(0)\beta\sum f]}\cdot\left\{\frac{2\beta}{3}\frac{1}{\sum f}\left[\sum\left(\frac{k_1^2}{2m}\right)f\right]\right\}^{-1}[(\omega+i\varepsilon)^2-\omega_0^2]^{-1};$$

$$(5.14b)\quad \omega_0^2 = \left\{\frac{k^2}{2m}\frac{1+2\beta v_d(0)\sum f}{1+\beta v_d(0)\sum f}\right\}\cdot\left\{\frac{2\beta}{3}\frac{1}{\sum f}\cdot\left[\sum\left(\frac{k_1^2}{2m}\right)f\right]\right\}^{-1} = k^2\,\frac{\chi_S^{-1}V}{m\langle N\rangle}\cdot\left\{\frac{2\beta}{3}\left(\frac{\overline{k_1^2}}{2m}\right)\right\}^{-1}.$$

Except for the « spurious » last factor in ω_0^2 and except for the fact that the isothermal compressibility appears in place of the isentropic function, this result is precisely (2.9). The spurious factor results from the fact that in (2.5) we admitted a general $\varepsilon^0_{k_1}$; it reduces to unity when $\varepsilon^0_{k_1} = k_1^2/2m$. On the other hand the presence of the isothermal compressibility is a serious flaw. Before proceeding to rectify this error, we note that the present results *are* in agreement with the observations of equations (2.9) through (2.11) with respect to the existence of a single resonance. Moreover, provided one recognizes that to the first order in v

$$(5.15)\quad \langle N\rangle = [1+\beta v_d(0)\sum f]^{-1}\cdot\sum f,$$

it is easy to verify that the expressions (5.14) are quantitatively self consistent.

The incorrect Newton resonance of the preceding paragraph was obtained by utilizing the conditions (5.3*a*) and (5.3*b*). If in addition we include the energy conservation condition (5.3*c*), we shall obtain the correct Laplace result. Specifically, imposing the three conditions (5.3), we find

$$(5.16)\quad B_1 = B_2 = B_{1a} = B_{1b} = B_{1c} = B_4 = 0; \qquad [N, H] = [\boldsymbol{P}, H] = [H, H] = 0.$$

Passing immediately to the long-wavelength low-density limit, we obtain in much the same way as before

$$(5.17b)\quad \mathscr{D}(\omega + i\varepsilon; k)_0 = -\frac{k^2 \sum f}{m[1 + v_d(0)\beta \sum f]} \cdot \frac{\omega + i\varepsilon}{\omega + i\varepsilon} \cdot \left\{\frac{2\beta}{3}\left(\frac{\overline{k_1^2}}{2m}\right)\right\}^{-1} \cdot [(\omega + i\varepsilon)^2 - \tilde{\omega}_0^2]^{-1},$$

$$(5.17b)\quad \tilde{\omega}_0^2 = \frac{k^2}{\beta m}\left\{\frac{1 + 2\beta v_d(0) \sum f}{1 + \beta v_d(0) \sum f}\right\} \cdot \left\{\frac{(\varepsilon - \bar{\varepsilon})^2[1 + 2\beta v_d(0) \sum f] + \beta^{-2}[1 + \beta v_d(0) \sum f]}{(\varepsilon - \bar{\varepsilon})^2[1 + 2\beta v_d(0) \sum f]}\right\} \left\{\frac{2\beta}{3}\left(\frac{\overline{k_1^2}}{2m}\right)\right\}^{-1} = \gamma\omega_0^2,$$

where

$$\bar{X} \equiv [\sum f(\varepsilon^0_{k_1}) X(k_1)]/[\sum f(\varepsilon^0_{k_1})]$$

and

$$(5.18)\quad \gamma = \frac{(\varepsilon - \bar{\varepsilon})^2[1 + 2\beta v_d(0) \sum f] + \beta^{-2}[1 + \beta v_d(0) \sum f]}{(\varepsilon - \bar{\varepsilon})^2[1 + 2\beta v_d(0) \sum f]}.$$

Expression (5.17) is easily recognized as the expected result. The ratio (5.18) may be easily verified for the limiting case $v \to 0$. In (5.17a) we have retained the trivial ratio $(\omega + i\varepsilon)/(\omega + i\varepsilon)$ to emphasize that in a certain qualitative sense there exists at this stage in the calculation a mechanism whereby the constant c of (2.10) may be different from zero. This constant, which arises from a branch line in the neighborhood of $\omega = 0$, could be determined directly from the integral $\int d\omega (\exp[\beta\omega] - 1)^{-1} A(\omega)$, if one were to retain in a calculation of the above type higher-order terms in the expansion (5.6).

By including such higher-order terms in ω/Γ, we would also find as before that there exists for the sound resonance a finite half-width on the order of $k^2(|\nabla\varepsilon|/\Gamma)_{\text{ave}}$. The « average » associated with this expression has an exceedingly complex structure which depends to a limited extent upon the terms of the expansion (5.2) which were neglected in (5.7). This last assertion may be easily verified if one repeats the calculation of this work, retaining several additional terms of the expansion (5.2) beyond (5.7). In much the same way, one may verify that the $\boldsymbol{k}$-dependence of $K_{12}(\,)_{\boldsymbol{k}}$ and $K_{34}(\,)_{\boldsymbol{k}}$ first appears in the neglected damping terms.

It is clear from these results that, provided the ratio $k(|\nabla\varepsilon|/\Gamma)_{\text{ave}}$ is sufficiently small, a dilute weakly interacting M-B gas will display a sound resonance at the frequency predicted by classical arguments. This resonance has its origin in the single-particle damping collisions and depends upon the conservation laws satisfied by such collisions. The shift from isothermal to isentropic compressibility induced by energy conservation is particularly noteworthy. As the wave number k increases or the effective interaction strength as measured by Γ decreases, the ratio $k(|\nabla\varepsilon|/\Gamma)_{\text{ave}}$ will gradually become nonnegligible. In that case the sound propagation properties will become altered.

By considering increasingly higher-order terms of the expansion (5.6), one can compute those changes in propagation characteristics by the method outlined above.

* * *

I am indebted to Prof. P. C. Martin, Prof. C. Bloch, Dr. C. de Dominicis, and Dr. L. Verlet for helpful discussions and their enlightening comments relative to many-particle problems. I wish also to express my appreciation for the hospitality extended to me in France by Prof. M. Lévy and the Laboratoire de Physique Théorique et Hautes Energies de L'Université de Paris.

Appendix

In this appendix we have tabulated the various functions $\{B, C\}$ mentioned n Sect. 5 and have indicated the explicit form of the matrix $[M]$. To simplify the presentation, we have suppressed co-ordinate indices. Except for the vector magnitude k and the frequency $\omega + i\varepsilon$ (represented simply as ω), all functions are implicitly functions of $k_1\xi_1$, the indices over which the indicated summation extends.

$$B_1 = B_{1a}B_{1b} - B_{1c}^2;$$

$$B_{1a} = \sum \frac{g_{01}g_{01}}{\Gamma f(1 \pm f)} - 1; \qquad B_{1b} = \sum \frac{g_{02}g_{02}}{\Gamma f(1 \pm f)} - 1; \qquad B_{1c} = \sum \frac{g_{01}g_{02}}{\Gamma f(1 \pm f)};$$

$$B_2 = \sum \frac{g_1 g_1}{3\Gamma f(1 \pm f)} - 1;$$

$$B_3 = \sum \frac{g_1 g_1}{3\Gamma^2 R_1 f(1 \pm f)};$$

$$B_4 = B_{1b}B_{4a} + B_{1a}B_{4b} - 2B_{1c}B_{4c};$$

$$B_{4a} = \sum \frac{g_{01}g_{01}}{\Gamma^2 R_1 f(1 \pm f)}; \qquad B_{4b} = \sum \frac{g_{02}g_{02}}{\Gamma^2 R_1 f(1 \pm f)}; \qquad B_{4c} = \sum \frac{g_{01}g_{02}}{\Gamma^2 R_1 f(1 \pm f)};$$

$$B_5 = B_{4a}B_{4b} - B_{4c}^2;$$

$$B_{6a} = \sum \frac{g_1 g_{01}|\nabla\varepsilon|}{\Gamma^2 R_1 f(1 \pm f)}; \qquad B_{6b} = \sum \frac{g_1 g_{02}|\nabla\varepsilon|}{\Gamma^2 R_1 f(1 \pm f)};$$

$$C_0 = \sum \frac{f(1 \pm f)|\nabla\varepsilon|^2}{\Gamma^2 R_1};$$

$$C_1 = \sum \frac{g_{01}}{\Gamma}; \qquad C_2 = \sum \frac{g_{02}}{\Gamma};$$

$$C_3 \quad \sum \frac{g_1|\nabla\varepsilon|}{\Gamma}.$$

$$
\begin{bmatrix}
\nu_{ss}(k)\cdot\sum\frac{-\beta k^2 f(1\pm f)|\nabla\varepsilon|^2(1+\dots)}{3\Gamma^2R_1}-1 & \sum\left\{1+i\frac{\omega}{R_1\Gamma}-\frac{3\omega^2+k^2|\nabla\varepsilon|^2}{3\Gamma^2R_1^2}+\dots\right\}\frac{g_{0s}}{\Gamma} & \sum\left\{1+i\frac{\omega}{R_1\Gamma}-\frac{3\omega^2+k^2|\nabla\varepsilon|^2}{3\Gamma^2R_1^2}+\dots\right\}\frac{g_{0r}}{\Gamma} & \sum\frac{-ik|\nabla\varepsilon|g_1}{9\Gamma^2R_1}\left\{1+i\frac{2\omega}{\Gamma R_1}+\dots\right\} \\
\nu_{sr}(k)\cdot\sum\frac{-\beta k^2 g_{0s}\ \nabla\varepsilon\ ^2(1+\dots)}{3\Gamma^2R_1} & \sum\left\{1+i\frac{\omega}{R_1\Gamma}-\frac{3\omega^2+k^2|\nabla\varepsilon|^2}{3\Gamma^2R_1^2}+\dots\right\}\frac{g_{0s}g_{0s}}{\Gamma f(1\pm f)}-1 & \sum\left\{1+i\frac{\omega}{R_1\Gamma}-\frac{3\omega^2+k^2|\nabla\varepsilon|^2}{3\Gamma^2R_1^2}+\dots\right\}\frac{g_{0s}g_{0r}}{\Gamma f(1\pm f)} & \sum\frac{-ik_1\nabla\varepsilon\ g_1g_{0s}}{9\Gamma^2R_1f(1\pm f)}\left\{1+i\frac{2\omega}{\Gamma R_1}+\dots\right\} \\
\nu_{rs}(k)\cdot\sum\frac{-\beta k^2 g_{0s}\ \nabla\varepsilon\ ^2(1+\dots)}{3\Gamma^2R_1} & \sum\left\{1+i\frac{\omega}{R_1\Gamma}-\frac{3\omega^2+k^2|\nabla\varepsilon|^2}{3\Gamma^2R_1^2}+\dots\right\}\frac{g_{0s}g_{0s}}{\Gamma f(1\pm f)} & \sum\left\{1+i\frac{\omega}{R_1\Gamma}-\frac{3\omega^2+k^2|\nabla\varepsilon|^2}{3\Gamma^2R_1^2}+\dots\right\}\frac{g_{0r}g_{0s}}{\Gamma f(1\pm f)}-1 & \sum\frac{-ik|\nabla\varepsilon|g_1g_{0s}}{9\Gamma^2R_1f(1\pm f)}\left\{1+i\frac{2\omega}{\Gamma R_1}+\dots\right\} \\
\nu_{ss}(k)\cdot\sum\frac{-ik\beta g_1|\nabla\varepsilon|}{\Gamma}\left\{1+i\frac{\omega}{\Gamma R_1}+\dots\right\} & \sum\frac{-ik|\nabla\varepsilon|g_1g_{0s}}{\Gamma^2R_1f(1\pm f)}\left\{1+i\frac{2\omega}{\Gamma R_1}+\dots\right\} & \sum\frac{-ik|\nabla\varepsilon|g_1g_{0s}}{\Gamma^2R_1f(1\pm f)}\left\{1+i\frac{2\omega}{\Gamma R_1}+\dots\right\} & \sum\frac{g_1g_1}{3f\Gamma(1\pm f)}\left\{1+i\frac{\omega}{\Gamma R_1}-\frac{\omega^2+\frac{2}{5}k^2|\nabla\varepsilon|^2}{\Gamma^2R_1^2}+\dots\right\}
\end{bmatrix}=
$$

SUPPLEMENTO AL VOLUME XVII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1960
3° Trimestre

Dislocation Relaxation at High Frequencies (*).

P. G. BORDONI

Istituto di Fisica Matematica dell'Università - Pisa
Istituto Nazionale di Ultracustica «O. M. Corbino» - Roma

(ricevuto il 29 Giugno 1960)

CONTENTS — **1.** Introduction. – I. **Evidence of a dislocation relaxation.** – **2.** Experimental results. – **3.** Theory of dislocation motion. – II. **Fundamental parameters of dislocation relaxation and influence of different treatments.** – **4.** Activation energy and characteristic frequency. – **5.** Relaxation spectrum. – **6.** Frequency relaxation. – **7.** Subsidiary dissipation peak. – **8.** Influence of anneal, cold-work, impurities, strain-amplitude and neutron irradiation. – **9.** Relaxation peak in alloys. – **10.** Conclusions.

LIST OF SYMBOLS

b Burger's vector of a dislocation, taken as the distance between closest rows of atoms.
c Velocity of elastic waves; c_e, c_t, c_l, values for extensional, transversal and longitudinal waves; also a constant.
d Logarithmic decrement.
e Logarithmic constant = 2.718... .
f Vibration frequency at the temperature T ; f_m, value at the temperature T_m of maximum energy dissipation; f_0, relaxed value of frequency at the temperature T; $f_{m,0}$, limiting value of f_m at very high temperatures.
h Planck's constant.
l Dislocation lenght.
k Boltzmann constant.
m Linear density of a dislocation.
t Time.

(*) Invited paper at the Seminar on «Resonance and Relaxation in Metals» of the American Society for Metals, Chicago, Ill., November 1959. Published by the permission of the A.S.M.

w Lenght of a kinked segment of a dislocation.
x Axis parallel to the dislocation line.
y Axis normal to the dislocation line.

E Energy per unit lenght which opposes the motion of a dislocation; E_0, constant term in the series expansion of $E(y)$.
F_1 Function of the ratio $(H_{cr}/2H_k)$ considered in Donth's theory.
H Computed value of the activation energy; H_M, H_W, H_S, values given by Mason, Weertman, Seeger's formulae; H_k, energy of a single kink; H_{cr}, critical energy for a couple of kinks.
M Elastic modulus; M_R, M_U, relaxed and unrelaxed values.
N_0 Volume density of dislocations.
Q Coefficient of resonance (quality factor) of a solid. Usually the inverse value Q^{-1} is taken as a measure of energy dissipation; Q_m^{-1}, value of Q^{-1} at the peak.
S Total relaxation strenght; S_e, S_l, values for extensional and longitudinal vibrations.
T Absolute temperature; T_m, temperature of maximum energy dissipation.
U Vibration energy.
W Experimental value of the activation energy for the relaxation effect; W', value for the Niblett and Wilks peak; $\overline{W}$, central value of a spectrum of activation energies.
Y Attenuation function for a single relaxation time; $\overline{Y}_\tau$, $\overline{Y}_W$, values computed for the centers of a time- or of an energy-spectrum.

α Attenuation coefficient of vibration amplitude with space.
β Temperature coefficient of frequency $= -\partial \log f/\partial T$.
γ Parameter characterizing the width of a Fuoss-Kirkwood spectrum.
$\delta(\tau)$ Density of relaxation strenght in a dislocation spectrum; δ_m, maximum value of $\delta(\tau)$.
Δ Increment of some physical quantity; ΔU, increment of vibration energy; ΔW, increment of activation energy.
ε Strain; ε_{el}, ε_{an}, elastic and anelastic components.
η Parameter characterizing the width of a rectangular spectrum $= \frac{1}{2} \log \tau_2/\tau_1$:
ϑ Debye's temperature.
λ Wavelenght.
μ Second Lamé's constant.
ν Computed frequency of dislocation motion; ν_0, limiting value of ν at very high temperatures.
ϱ Volume density of a solid.
σ Stress; σ_p^0, Peierls' stress when thermal and quantum mechanical fluctuations are disregarded.
τ Characteristic time of a relaxation effect, at the temperature T; τ_0, limiting value of τ at very high temperatures; $\overline{\tau}$ central value of a relaxation spectrum; τ_1, τ_2, values at which the density $\delta(\tau)$ is $1/\sqrt{2}$ of its maximum value δ_m; $\overline{\tau}_0$ limiting value of $\overline{\tau}$ at very high temperatures.

1. – Introduction.

The first experimental evidence of a relaxation effect due to the dislocations was found a few years ago in some low-temperature measurements of the

dissipation of elastic energy in lead, copper, aluminium and silver [A-1] (1). An exhaustive investigation of the effect in copper employing both polycrystalline specimens and single crystals, was successively made by NIBLETT and WILKS [A-4, A-8], by CASWELL [A-11] by PARÉ [A-13] by THOMPSON, GLASS and HOLMES [A-14, A-17], by NUOVO, VERDINI and the A. [A-15, A-16]. Measurements were also made on aluminium, by HUTCHINSON, HUTTON and FILMER [A-5, A-6, A-12] and by EINSPRUCH and TRUELL [A-9]; on lead by BOEMMEL and MASON [A-2, A-7]. The same effect was recently found in gold, palladium and platinum by NUOVO, VERDINI and the A. [A-18], and in gold-silver alloys by the same AA. and BARDUCCI [A-19].

The first attempt to give a theorethical model for the dislocation motions to which the effect is related, was made by MASON [B-2, B-4] assuming that a dislocation line between two pinning points could be removed from its potential well, and that the stresses due to the elastic waves would make the potential well asymmetrical, giving rise to a macroscopic anelastic strain and producing a relaxation effect. This model lays perhaps too much emphasis on the effect of impurity content, and was discussed by WEERTMAN [B-3, B-5] and by SEEGER [B-8], who suggested that the dislocations may be confined to certain crystallographic directions by the Peierls' stress, and may form pair of kinks, under the combined action of the thermal fluctuations and of the applied stress. The energy required by the formation of a kink pair is an intrinsic property of dislocations, and this removes the main difficulty of MASON's model. The computation of the fundamental parameters associated with the relaxation effect has been successively improved by SEEGER, DONTH and PFAFF [B-9, B-10], who have shown that the relation between the relaxation frequency and the temperature does not require any special hypothesis, but is derived from the theory of stochastic processes.

The present review of the actual knowledge on dislocation relaxation is divided in two parts. In the first the experimental evidence of a thermally activated relaxation effect due to the dislocations is discussed, together with the theoretical explanations which have been proposed for this effect.

In the second part, all the basic parameters associated with the effect are considered in detail, and their experimental values are compared with those given by the theory. The influence of thermal and mechanical treatments, of impurity content, of high energy irradiation and of strain amplitude are also examined from a quantitative standpoint. Unsolved or unexplored sides of dislocation relaxation are brought to view by this analysis and hints are derived for future investigation.

---

(1) The Section A of the bibliography lists the esperimental contributions; the theoretical papers are collected in Section B.

PART I. – Evidence of a dislocation relaxation.

2. – Experimental results.

When the dissipation coefficient Q^{-1} is measured as a function of the temperature for a resonant specimen of a f.c.c. metal (²) a peak is generally

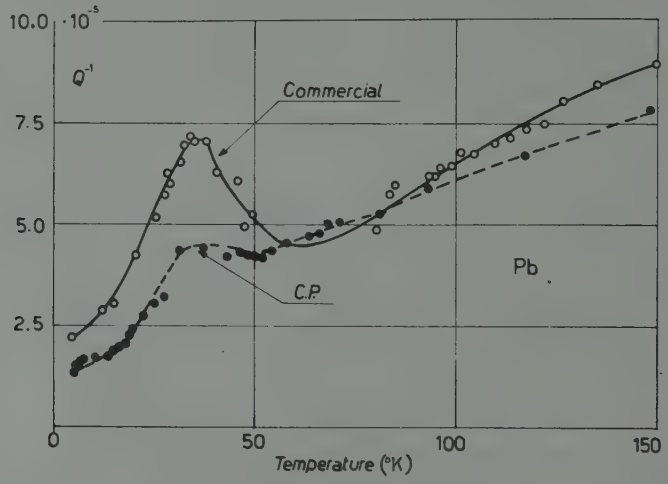


Fig. 1. – Relaxation peak in commercial and chemically pure polycrystalline lead: vibration frequency $\simeq$ 10 kHz [A-1].

(²) The anelastic behaviour of a solid may be evaluated by dinamic measurements either on standing waves or on travelling waves. The coefficient Q^{-1} obtained in the first case is proportional to the ratio of the energy ΔU dissipated in half a cycle to the stored energy U

$$Q^{-1} = \pi^{-1} \frac{\Delta U}{U} .$$

The coefficient Q^{-1} gives also the tangent of the phase angle between a sinusoidal stress and the corresponding strain.

Experiments on travelling waves of wave lenght λ give the attenuation coefficient α which measures the decrease of vibration amplitude with space. Both types of measurements are equivalent, as far as the anelastic behaviour of solids is concerned, the parameters Q^{-1} and α being related by the equation

$$Q^{-1} = \alpha \lambda \pi^{-1} .$$

The largest number of data on dislocation relaxation has been obtained by measurements on resonant systems: hence Q^{-1} will be taken in the following as a measure of the anelastic behaviour.

found in the experimental curve at a temperature which is considerably lower than the room temperature when the vibration frequency is in the acoustic or ultrasonic range. The temperature T_m and the height of the peak Q_m^{-1} may differ considerably from one metal to another, even when the vibration frequencies have almost the same value. For instance in lead and copper for a vibration frequency near to 10 kHz the temperature T_m has the value 35 °K for the first, and 82 °K for the second (Figs. 1, 2). The difference between the values of Q_m^{-1} for the same metals is even larger, as the height of the peak is about $7 \cdot 10^{-5}$ for the former, whilst it exceeds $400 \cdot 10^{-5}$ for the latter. Differences of the same order are found between the attenuation peak observed in aluminium (Fig. 3) for a vibration frequency near 40 kHz and the peaks of silver, gold, palladium and platinum (Fig. 4).

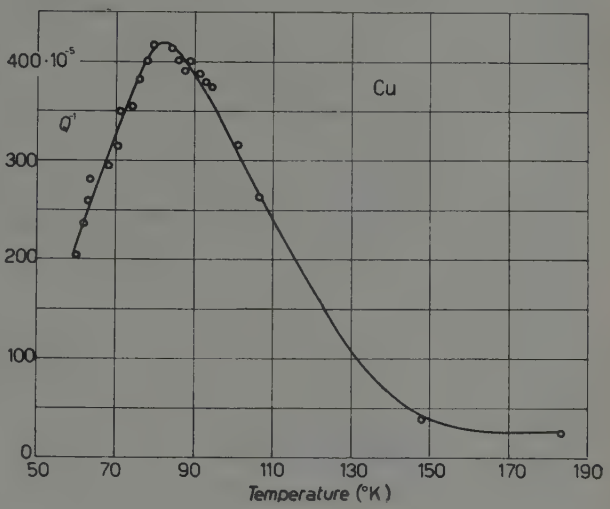


Fig. 2. – Relaxation peak in technically pure polycrystalline copper: vibration frequency $\simeq$ 10 kHz [A-16].

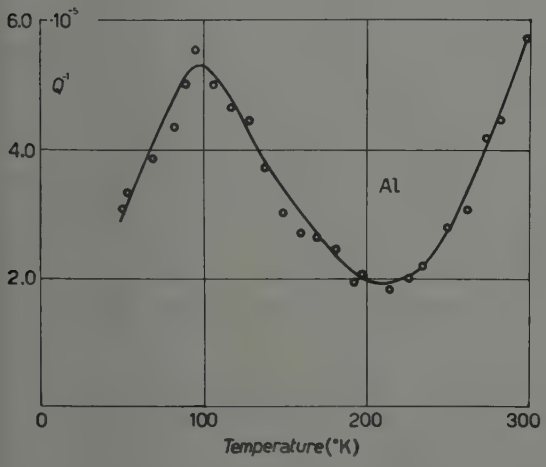


Fig. 3. – Relaxation peak in chemically pure polycrystalline aluminium (99.9%): vibration frequency $\simeq$ 40 kHz [A-1].

A common characteristic of all the above peaks is that they are a very stable feature of the low temperature dissipation measurements, no appreciable differences being found between measurements in successive runs. Moreover the temperature T_m corresponding to each vibration frequency has a characteristic value for each metal, which is little or not affected by thermal and mechanical treatments, by impurity content, by high energy irradiation and by vibration amplitude, as it will

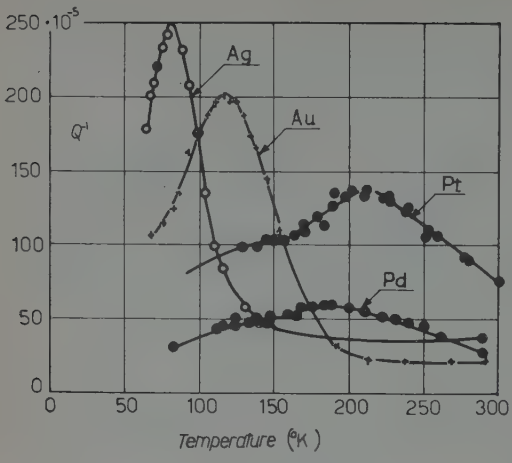


Fig. 4. – Relaxation peak in polycrystalline specimens of chemically pure silver (99.80%), $f \simeq 50$ kHz; gold (99.88%), $f \simeq 16$ kzH; palladium (99.91 %), $f \simeq 88$ kHz; platinum (99.89%), $f \simeq 112$ kHz [A-18].

be shown later. It must also be noticed that for copper, silver and gold the peaks are very high, Q_m^{-1} being many times larger than the energy dissipation measured at room temperature.

From the first experiments on lead it was evident that when measurements are made at different frequencies, the temperature of the dissipation peak increases with frequency (Fig. 5). This result was confirmed by BOEMMEL's measurements on the same material [A-2] and by all the measurements made on the other metals in which the effect has been found, as it is clearly shown by Fig. 6 in which are collected the experimental curves obtained for copper in a wide frequency range (1.8 kHz ÷ 6.4 MHz). The measurements are accurate enough to show that the logarithm of the vibration frequency f_m is proportional to the inverse temperature of the peak T_m

$$\ln f_m = -\frac{c_1}{T_m} + c_2 \,, \tag{1}$$

where c_1 and c_2 are two positive constants.

As it was pointed out by MASON [A-2, A-4] this frequency dependence of the temperature of maximum dissipation indicates a thermally

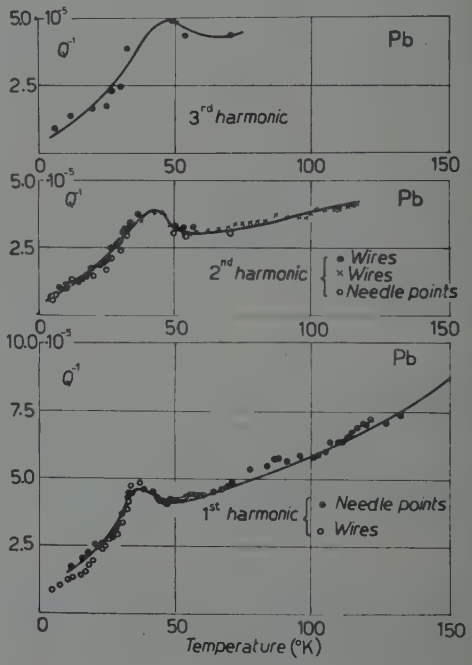


Fig. 5. – Frequency dependence of the temperature of maximum dissipation in chemically pure polycrystalline lead: frequency of the first harmonic $\simeq 10$ kHz [A-1].

activated relaxation effect, associated with a characteristic time τ which depends upon the temperature according to an Arrhenius equation

$$\tau = \tau_0 \exp\left[\frac{W}{kT}\right], \tag{2}$$

where W is the activation energy and τ_0 gives the limiting value of the characteristic time at very high temperatures. This hypothesis explains not only

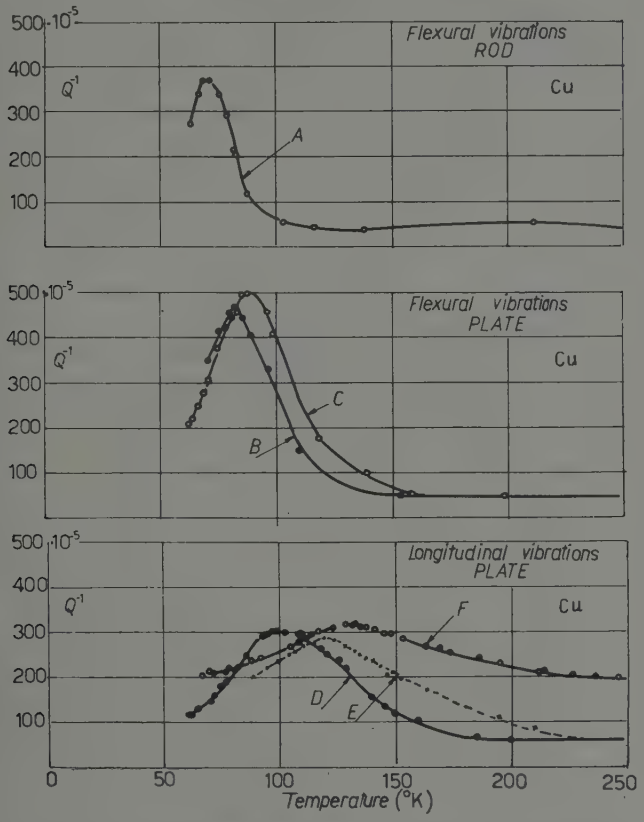


Fig. 6. – Frequency dependence of the temperature of maximum dissipation in technically pure polycrystalline copper. Curve A, frequency 1.8 kHz; B, 13 kHz; C, 45 kHz; D, 550 kHz; E, 2.7 MHz; F, 6.4 MHz [A-16].

the observed dependence of Q^{-1} upon frequency and temperature, but also the temperature dependence of the resonant frequency.

In fact the theory of relaxation effects shows that the energy dissipation depends upon T and f through the product $f \cdot \tau$ according to the equation

$$Q^{-1} = Q_m^{-1} \operatorname{sech}\left[\ln\left(2\pi f\tau\right)\right] = Q_m^{-1} \operatorname{sech}\left[\ln\left(2\pi f\tau_0 \exp\left[\frac{W}{kT}\right]\right)\right]. \tag{3}$$

The dissipation reaches its maximum value when $2\pi f\tau = 1$. Then for a given vibration frequency f_m, the corresponding temperature T_m of the peak is obtained multiplying (2) by $2\pi f_m$, putting T_m instead of T and equating to unity

$$\ln f_m = -\frac{W}{kT_m} - \ln (2\pi\tau_0) \,. \tag{4}$$

Equation (4) is the same as the experimental relation (1), provided the proper values are taken for the activation energy and the limiting time

$$\begin{cases} W = kc_1 \\ \tau_0 = (2\pi)^{-1} \exp [-c_2] \,. \end{cases} \tag{5}$$

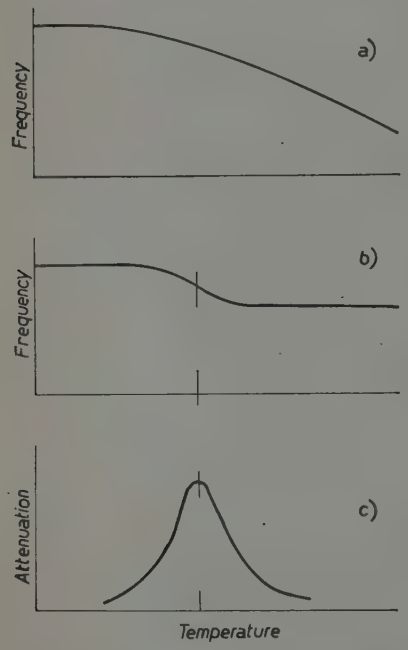


Fig. 7. – *a*) Temperature dependence of frequency due to thermal dilatation. *b*) Temperature dependence of frequency for a thermally activated relaxation effect. *c*) Temperature dependence of dissipation for a thermally activated relaxation effect.

The relaxation theory predicts also that the resonant frequency of a given vibration mode of the specimen must change with temperature according to the relation

$$\frac{f - f_0}{f_0} = \frac{Q_m^{-1}}{2} \left\{ 1 + \mathrm{tgh} \left(\ln 2\pi f \tau_0 \exp \left[\frac{W}{kT} \right] \right) \right\} , \tag{6}$$

where f_0 is the relaxed value of the resonant frequency. Equation (6) does not account for the monotonic frequency changes due to the effect of thermal expansion on the size of the specimen and on the elastic moduli. When the latter variation is added to that represented by (6), an inflexion point must be found in the frequency-temperature curve at a temperature very near to T_m as it may be seen by Fig. 7, adding the values of frequency given by the curves *a*) and *b*).

In accordance with the theory, the experimental frequency-temperature curves exhibit an inflexion point corresponding to the temperature of maximum damping. As it is required by (6) the inflexion is more pronounced

in those metals which have a larger value of Q_m^{-1}, as it is the case with copper (Fig. 8).

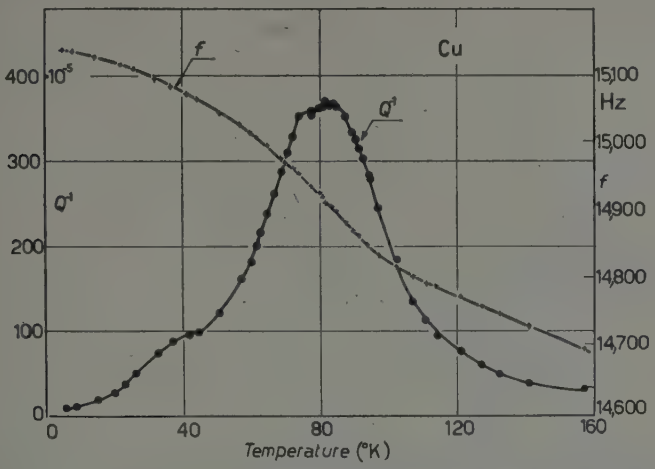


Fig. 8. – Inflexion point in the frequency-temperature curve at the temperature of maximum dissipation for technically pure polycrystalline copper [A-16].

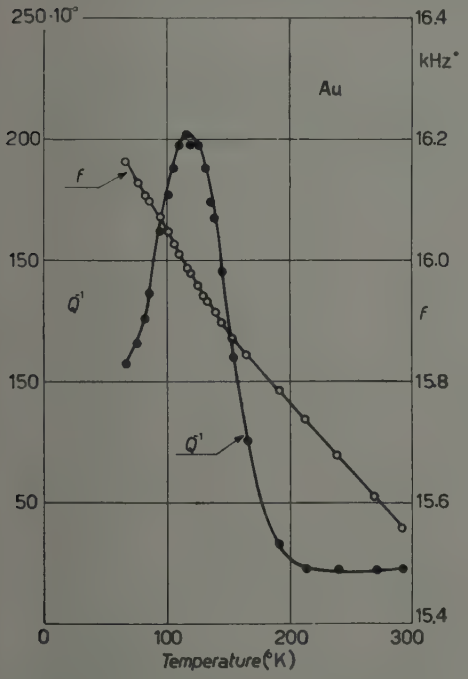


Fig. 9. – Inflexion point in the frequency-temperature curve at the temperature of maximum dissipation for chemically pure (99.88%) polycrystalline gold [A-20].

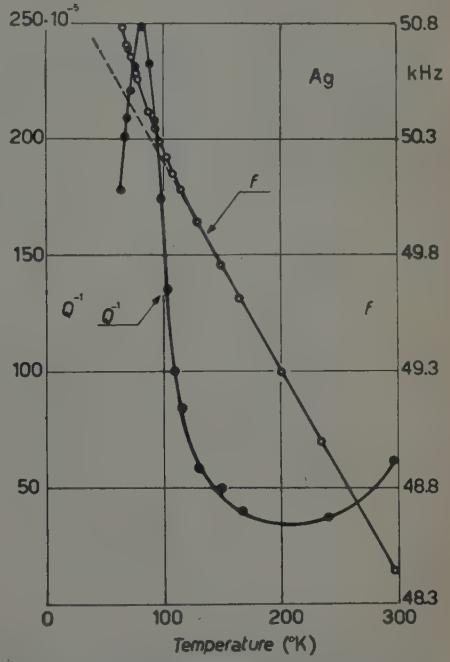


Fig. 10. – Inflexion point in the frequency-temperature curve at the temperature of maximum dissipation for chemically pure (99.80%) polycrystalline silver [A-20].

The inflexion is clearly seen also in gold (Fig. 9) and in silver (Fig. 10). notwithstanding the limited temperature range covered by the experimental

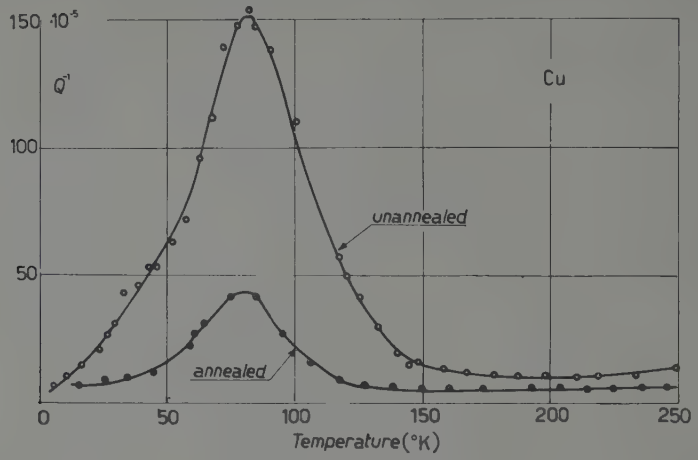


Fig. 11. – Reduction of the relaxation peak due to anneal; influence of a 2-hours treatment at 175 °C on commercial polycrystalline copper: vibration frequency $\simeq$ 30 kHz [A-1].

data. In palladium and platinum, whose dissipation peaks are smaller and flatter (Fig. 4) the inflexion is less evident, whilst in lead and silver the peaks are so small (Figs. 1, 3) that the measurements are not accurate enough to show a frequency relaxation which does not exceed 10^{-4}.

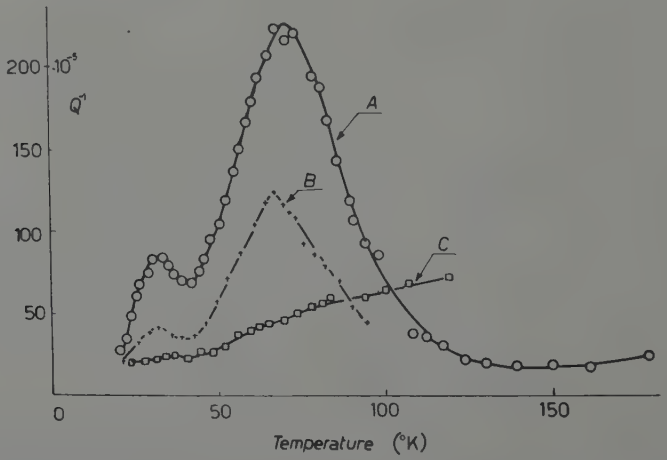


Fig. 12. – Dependence of he height of the relaxation peak upon the annealing temperature in high purity (99.999%) oxygen-free polycrystalline copper. Curve *A*, after straining 8.4%; *B*, after 1 h anneal at 180 °C; *C*, after 1 h anneal at 350 °C; vibration frequency $\simeq$ 1 kHz [A-8].

To conclude the general discussion of the experimental results, it must be shown that the relaxation effect is due to the dislocations. To this purpose the influence of thermal and mechanical treatments must be considered

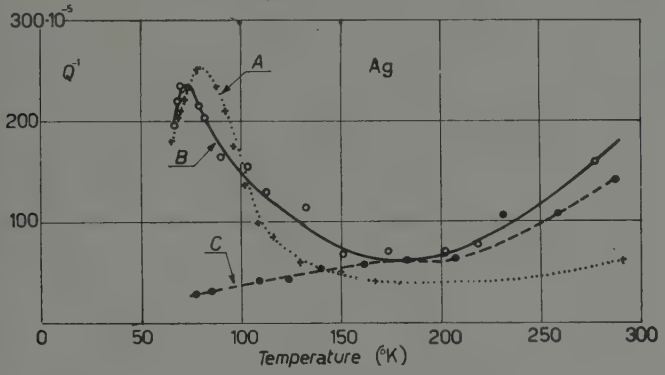


Fig. 13. – Dependence of the height of the relaxation peak upon the annealing temperature in chemically pure (99.80%) polycrystalline silver. Curve *A*, before anneal; *B*, after a 4-hours treatment at 175 °C; *C*, after an additional 5-hours treatment at 225 °C; vibration frequency $\simeq$ 56 kHz [A-20]

together with the measurements which have been made on single crystals. It has been found that any treatment which reduces the number of dislocations, such as anneal, decreases the height of the peak, as it shown in Fig. 11 for a not very pure copper specimen. The reduction of Q_m^{-1} depends upon the annealing temperature: if this temperature is high enough, the peak may be cancelled, as it is shown by Figs. 12 and 13 for copper and silver. The effect of anneal can be followed through the time-changes of frequency at constant temperature when the final temperature of the treatment has been reached (Fig. 14). The increase of frequency shows a corresponding decrease

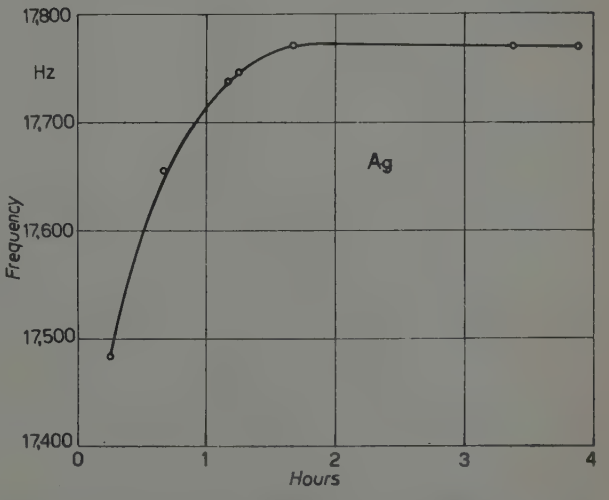


Fig. 14. – Influence of anneal on polycrystalline silver (99.80%). Increase of the resonant frequency with time at the constant temperature of 225 °C [A-20].

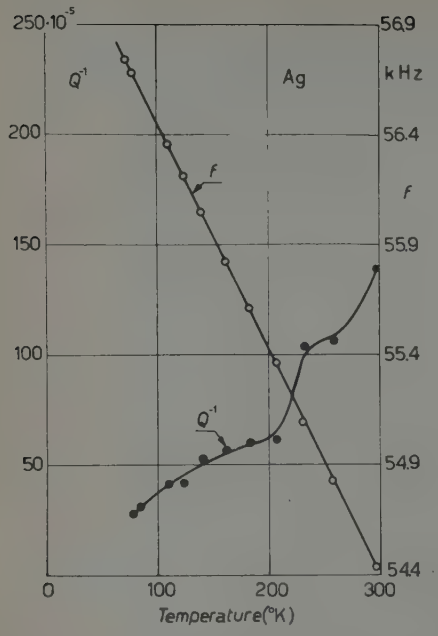


Fig. 15. – Linear dependence of the resonant frequency upon temperature in polycrystalline silver (99.80%), after a 5 hours anneal at 225 °C [A-20].

of the number of dislocations active in the relaxation process, and explains the reduction found in the low-temperature peak. After a complete anneal, when the peak has been cancelled, the frequency-temperature curve no longer exhibits any trace of the inflection due to the relaxation effect (fig. 15).

On the other hand, any treatment which increases the number of dislocations, such as cold work, increases also the height of the peak, when this is already present (Fig. 16), or reintroduces it, if the material has previously undergone a complete anneal (Figs. 17, 18). It may be added that the annealing temperature required to cancel the peak is considerably higher than room temperature, for all metals in which the effect has been found. As it was pointed out by SEEGER [B-8], this excludes that the relaxation can be due to the re-orientation of divacancies or similar defects created by the plastic deformation which anneal out rapidly at room temperature, and suggests that the peaks are produced by the motion of dislocation segments.

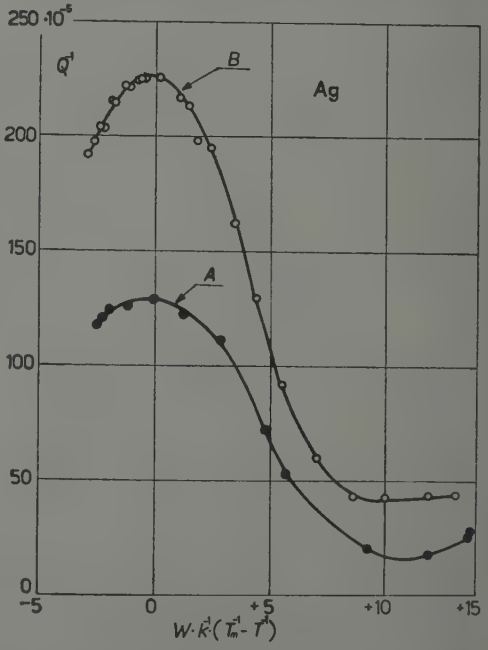


Fig. 16. – Increase of the height of the peak with the amount of cold work in polycrystalline silver (99.80%). Curve *A*, after a moderate cold rolling; curve *B*, after a further cold rolling to $\frac{2}{3}$ of previous thickness; vibration frequency $\simeq$ 56 kHz [A-20].

All the previous measurements were made on polycrystalline specimens. However the grain structure cannot be responsible for the dissipation peaks, owing to their low temperature. The experimental proof of this has been

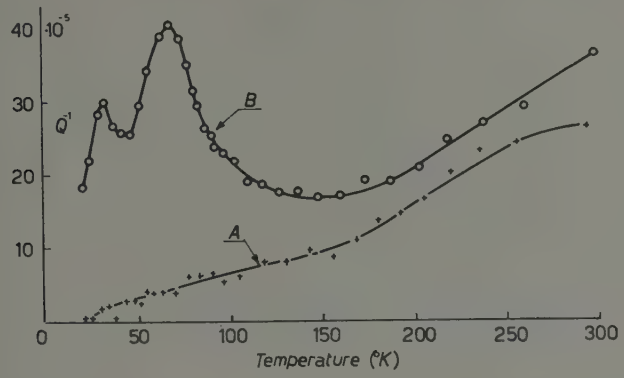


Fig. 17. – Increase of the height of the relaxation peak with the amount of cold work in polycrystalline copper (99.97%). Curve *A*, strained 0.17%; *B*, strained 5.5%; vibration frequency $\simeq$ 0.4 kHz [A-8].

obtained by the A. [A-1] by CASWELL [A-11] and by PARÉ [A-13] who have found the dissipation peak in copper single crystals. Within the limits of experimental accuracy, the temperature T_m corresponding to a given fre-

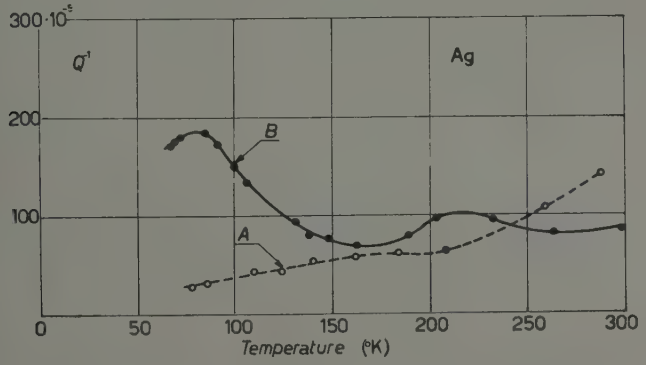


Fig. 18. – Increase of the height of the relaxation peak with the amount of cold work in polycrystalline silver (99.80%). Curve *A*, after a 5 hours anneal at 225 °C; *B*, after 2.7% permanent strain; vibration frequency $\simeq$ 56 kHz [A-20].

quency f_m is the same for polycrystalline specimens and single crystals and the effect of annealing treatments and cold work is also qualitatively the same for both types of materials (Fig. 19). This shows that the cause of the

dissipation maximum must be directly related to the geometrical imperfections of the crystal lattice and not to the grain structure.

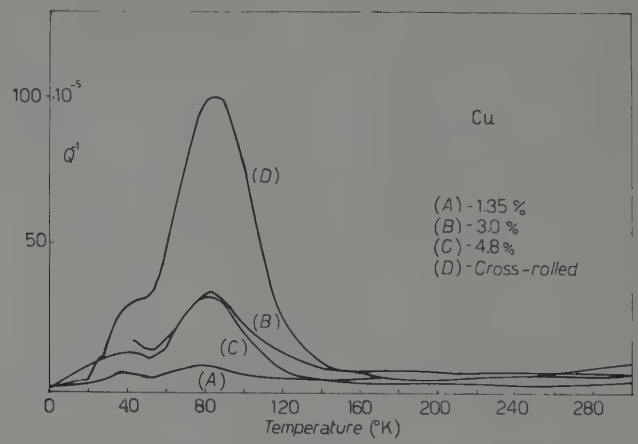


Fig. 19. – Relaxation peak in copper single crystals (99.999%) for various amounts of cold work. Vibration frequency $\simeq 40$ kHz [A-11].

3. – Theory of dislocation motion.

In order to explain how the dislocations can give rise to a thermally activated relaxation effect, some model must be found for their motion under the combined action of thermal fluctuations and mechanical stress. The

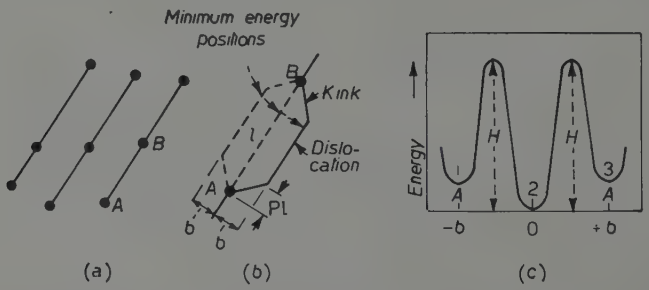


Fig. 20. – Mechanism for the dislocation motion according to Mason [B-2, B-4].

next step is to relate the motion of all the dislocations with the formation of an anelastic strain. The formal theory of relaxation can be successively applied to this strain to find the frequency- and temperature-dependence of the energy dissipation.

In the model proposed by Mason [B-2, B-4] a dislocation line is pinned

at points A, B by impurity atoms (Fig. 20*a*). The lowest-energy position corresponds to a straight segment, but other equilibrium positions may exist of somewhat higher energy, in which the dislocation is *kinked* (Fig. 20*b*). The central part of the dislocation is then displaced in the glide plane of an amount b, equal to the distance between closest lines of atoms in the same plane. Owing to the Peierls' stress, a dislocation must overcome a potential barrier H to go into the nearest equilibrium position (Fig. 20*c*). To compute this barrier, MASON assumed that the whole dislocation segment between the pinning points moves together. If the x axis is taken parallel to the dislocation line, and y normal to it, the potential per unit lenght $E(y)$ which opposes the motion of a dislocation may be represented between $y = -b/2$ and $y = +b/2$, by

$$E(y) = E_0 - \frac{\sigma_p^0 b^2}{2\pi} \cos \frac{2\pi y}{b}, \tag{7}$$

where $\sigma_p^0 = 1/b(\partial E/\partial y)_{\max}$, is the Peierls' stress required to force the dislocation through the steepest part of the potential, when thermal and quantum mechanical fluctuations are disregarded. Neglecting the changes in dislocation lenght, the height H_M of the potential barrier for this model, which is given by the maximum value of $l \cdot [E(y) - E(0)]$, that is

$$H_M = \frac{\sigma_p^0 b^2 l}{\pi}, \tag{8}$$

is proportional to the dislocation lenght l. This would lead to a strong dependence of the temperature of the peak on the parameters which control l, that is the dislocation density and the impurity content. As no such dependence is observed experimentally, some modification must be introduced in MASON's model, considering types of dislocation motions not controlled by the dislocation lenght l.

As it has been pointed out by WEERTMAN [B-5] it is not necessary for the whole dislocation segment to move at once in going from

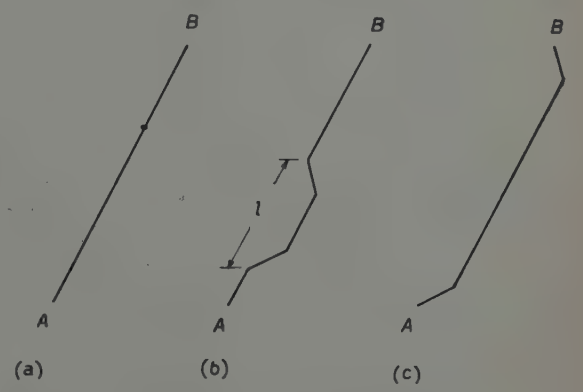


Fig. 21. – Mechanism for the dislocation motion according to WEERTMAN [B-3, B-5] and to SEEGER [B-8, B-10].

one equilibrium position to another. A small kinked segment of lenght $l' < l$ may move over the potential barrier, a lower limit for l' being given by the lenght $2w$ of a kink pair (Fig. 21). Taking MASONS's estimate for w

$$w = b\left(\frac{\mu}{\sigma_p^0}\right)^{\frac{1}{2}}, \tag{9}$$

and inserting $2w$ in equation (8) in the place of l, an approximate value is found for the energy H_w required by the formation of a kink pair

$$H_w = \frac{2}{\pi}\mu b^3\left(\frac{\sigma_p^0}{\mu}\right)^{\frac{1}{2}}, \tag{10}$$

where μ = shear modulus.

According to equation (10), the formation of a kink pair has an « intrinsic » character, its energy being independent of the particular lenght of dislocation lines and of their interaction with impurity atoms. Hence this model is better suited to explain the relaxation effect then the « rigid motion » model associated with the energy (8). However the derivation of (10) is somewhat lacking in mathematical rigour, as the maximum value of potential (7) is applied to a kinked line which is not parallel to the direction of maximum energy; moreover the approximate value (9) is taken as the lenght of a kink, and the changes in dislocation lenght are neglected.

A more satisfactory computation of the energy associated with the formation of kinks has been given by SEEGER [B-8]. The starting point is the differential equation for the shape of an unpinned dislocation line, whose minimum energy position is parallel to the x axis

$$E(y)\frac{\partial^2 y}{\partial x^2} = -\frac{\mathrm{d}E(y)}{\mathrm{d}y} + b\sigma + m\frac{\partial^2 y}{\partial t^2}, \tag{11}$$

where σ = resolved shear stress in the glide system; m = density per unit lenght. The periodic potential per unit lenght $E(y)$ may be represented by (7) for any value of y, a further refinement in the representation of $E(y)$ being not significant in the present state of theory (3). As $\sigma_p^0 b^2/2\pi \ll E_0$, the latter value will be substituted to $E(y)$ on the left side of (11). Then in the static case, with no external stresses ($\sigma = 0$) the equation becomes

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = \frac{\sigma_p^0 b}{E_0}\sin\frac{2\pi y}{b}. \tag{12}$$

(3) All the potential minima have now the same value, the dislocation being *unpinned*. The period has been taken equal to b.

A solution of (12) representing a single kink of amplitude b in the positive direction of the y axis is given by

$$y(x) = \frac{2b}{\pi} \operatorname{tg}^{-1} \left\{ \exp \left[x \left(\frac{2\pi\sigma_p^0}{E_0} \right)^{\frac{1}{2}} \right] \right\}. \tag{13}$$

The displacement y tends asymptotically to zero when x tends to $-\infty$ and to b when x tend to $+\infty$. Comparing (13) with the estimate (9) for the kink width w, it is found that the latter corresponds to a dislocation displacement differing by less then $0.04b$ from its final value ([4]).

A straight dislocation line lying along one of the lines $y = nb$ ($n = 0, 1, 2, 3 \ldots$) has the energy E_0 per unit lenght. When the line is no longer straight, the energy is increased owing to the contemporary change in lenght and to the presence of the sinusoidal term in equation (7). If the angles made by the dislocation with the x-axis are small, the first part of the energy increase is proportional to $(\mathrm{d}y/\mathrm{d}x)^2$ while the second part is proportional to $(1 - \cos 2\pi y/b)$ The energy H_k of a single kink may then be computed inserting (13) into the expression

$$\int_{-\infty}^{+\infty} \left\{ \frac{1}{2} E_0 \left(\frac{\mathrm{d}y}{\mathrm{d}x} \right)^2 + \frac{\sigma_p^0 b^2}{2\pi} \left(1 - \cos \frac{2\pi y}{b} \right) \right\} \mathrm{d}x \,. \tag{14}$$

This gives for H_k ([5])

$$H_k = \frac{2}{\pi} \sqrt{\frac{2}{5\pi}} \cdot \mu b^3 \left(\frac{\sigma_p^0}{\mu} \right)^{\frac{1}{2}}. \tag{15}$$

To compute the energy H_s associated with the formation of a kink pair SEEGER assumes that the energies of two half-kinks corresponding to a displacement $y = b/2$, add approximately to the energy H_k of one complete kink, which is given by (15). This amounts to say that the position of maximum energy through which a dislocation segment has to pass when forming a pair of kinks is approximately reached when the central part of the kink pair is on the crest of the potential barrier, midway between two equilibrium positions. Moreover there is an attractive force between the kink pairs, which is particularly strong for small separations d of the kinks. This force is due to the fact the total lenght of the kinked dislocation diminishes if the kinks come together by sideways motion, and eventually annihilate each other. An opposite action is exerted by the applied shear stress; if σ is the resolved

([4]) It must be remembered that $\sigma_p^0 b^2/2\pi \ll E_0$.

([5]) Following SEEGER [B-8] the value $\mu b^2/5$ is taken for E_0.

value in the glide system, it exerts a force $\sigma \cdot b$ per unit lenght of the dislocation which tends to increase the slipped area in the glide plane on one side of the dislocation line. For every value of the applied stress, there is a critical separation d_{cr} corresponding to an unstable equilibrium of the kink pair under the combined action of the applied stress and of the mutual attraction. The energy associated with the attraction computed for $d = d_{cr}$ must be added to H_k to obtain the total energy barrier H_s which opposes the formation of a kink pair. An approximate computation of d_{cr} gives for the total energy [B-8]

$$H_s = H_k \left\{ 1 + \frac{1}{4} \ln \left(\frac{16 \sigma_p^0}{\pi \sigma} \right) \right\}. \tag{16}$$

A basic difference between the above equation and the expressions (8) and (10) is that the energy barrier given by (16) depends upon the applied stress. However σ enters the energy H_s only in a logarithmic term, and the strain value for most of the measurements is from 10^{-8} to 10^{-6}. The ratio σ/μ is of the same order as the strain, while the ratio σ_p^0/μ associated with relaxation measurements is generally found to be near $5 \cdot 10^{-4}$. The value of the last term in (16) is from 2.0 to 3.0 and a value 2.5 may be taken as a satisfactory average for most experimental conditions [B-11]. This gives for H_s

$$H_s = 3.5 H_k = 1.25 \frac{2}{\pi} \mu b^3 \left(\frac{\sigma_p^0}{\mu} \right)^{\frac{1}{2}}, \tag{17}$$

which differs only by the factor 1.25 from the value (10) given by WEERTMAN (6). It may be observed that H_s tends to infinity when the applied stress tends to zero. This does not mean that kink pairs are not formed when there is no applied stress, but that they annihilate very soon after their formation.

To compute the frequency ν of formation of kink pairs as a function of the absolute temperature T is a rather difficult problem. All the above theories avoid this difficulty *assuming* that ν may be represented by an Arrhenius equation

$$\nu = \nu_0 \cdot \exp \left[- \frac{H}{kT} \right], \tag{18}$$

where H is the energy of the potential barrier opposing the dislocation motion given by (8), (10) or (17) according to the model considered and to the anal-

(6) In Mason's treatment of Seeger's theory the factor 1.25 of eq. (17) is replaced by 2.0 as he takes the value $\mu b^2/2$ for E_0, instead of $\mu b^2/5$ as it is done by SEEGER; see [B-11], p. 270, eq. (9.61).

ysis of its motion; ν_0 is the vibration frequency of a straigh rigid dislocation line, near its position of minimum energy.

The value of ν_0 is immediately obtained for small oscillations ($y \ll b/2$) from equation (11) putting $\partial^2 y/\partial x^2 = 0$ and $\sigma = 0$, and substituting to the sinusoidal function its argument

$$m \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 2\pi\sigma_p^0 y = 0 \,. \tag{19}$$

Taking ϱb^2 for the linear density m, as suggested by ESHELBY [7] (ϱ = volume density) and introducing the velocity $c_t = (\mu/\varrho)^{\frac{1}{2}}$ of shear waves [8]

$$\nu_0 = \frac{c}{b\sqrt{2\pi}} \left(\frac{\sigma_p^0}{\mu}\right)^{\frac{1}{2}} . \tag{20}$$

The frequency ν given by (18) and (20) must be interpreted as the average number of kink pairs which are formed by a single dislocation in a time interval much larger than ν^{-1}.

When no applied stress is present, on the average the same number of kink pairs will be formed on both sides of a dislocation line in its glide plane. The average dislocation position coincides then with the straight equilibrium position of minimum energy, and no macroscopic strain is originated. An applied shear stress, much smaller than Peierls' stress, acting in the glide system of the dislocation, will cause the kinks to get further apart on one side, and nearer to each other on the opposite side, making their statistical distribution no longer symmetrical with respect to the dislocation line. The average position of the dislocation with respect to time is no more a straight line, but it is displaced towards the side in which the applied stress favors the production of kinks. A macroscopic strain is then produced by the asymmetry of the average dislocation motion. This strain is certainly not of an *elastic* type, as a finite time of the order of ν^{-1} is required by the dislocations to approach a new statistical distribution after a stress is suddenly applied.

The above remarks are of a purely qualitative type. The *anelastic* character of the strain due to the dislocations could be established on a quantitative basis showing its linear dependence on the applied stress and its exponential dependence on time, after a stress is suddenly applied. Both properties are sum-

---

(7) See J. D. ESHELBY: *Proc. Roy. Soc. London*, A **197**, 396 (1949).

(8) SEEGER [B-8], p. 660 (17), takes $(E_0/m)^{\frac{1}{2}}$ for c_t. On the other hand he gives the estimate $\mu b^2/5$ for E_0. Then his expression for ν_0 differs by a factor $\sqrt{5}$ from eq. (20). The latter coincides with Mason's value [B-11], p. 271.

marized by a stress-strain relation of the type [9]

$$\varepsilon(t) = \frac{\bar{\sigma}}{M}\left\{1 + \frac{\varepsilon_{an}}{\varepsilon_{el}}\left[1 - \exp\left[-\frac{t}{\tau}\right]\right]\right\}, \qquad \text{for } t > 0, \tag{21}$$

where $\varepsilon(t)$ is the total strain corresponding to a suddenly applied stress $\bar{\sigma}$; M is an elastic modulus; ε_{el}, ε_{an} are the elastic and anelastic fraction of the strain. No attempt is made in SEEGER's theory to derive equation (21) from the analysis of the model of dislocation motion. Such an attempt could eventually be made considering the average dislocation motion and the average equilibrium position under an applied stress. It is obvious that the formation of kink pairs is symmetrical with respect to such a position, owing to its *average* character. This means that the effect of applied stress which favors the production of kinks on one side is counteracted near the dislocations by a distribution of *elastic* stresses due to their average displacement which must therefore be proportional to the applied stress. The exponential dependence upon time shown by (21) can also be explained observing that the average dislocations motion towards a new equilibrium position, has an entirely statistical character, being controlled by the thermal fluctuations. The average motion is therefore irreversible, and the number of dislocations which in a given time make a given step towards the new equilibrium position must necessarily be a fixed fraction of the total number of dislocations which have not yet made the same step. Hence the strain-time relation is of an exponential type, as required by (21).

Once a relation of the type (21) has been proved or adopted, the formal theory of thermally activated relaxation effects shows that the temperature dependence of the dissipation and of the resonant frequency is represented by equations (3) and (6), and the theoretical treatment of the effect is complete from a semiquantitative standpoint. A further analysis of the anelastic strain due to the dislocations motion is however needed to find a relation between the potential barrier H (given by equations (8), (10) and (17)) and the activation energy W of equation (2). A similar relation is also needed between the relaxation time τ_0 considered in the same equation and the frequency ν_0 of dislocation motion, computed according to (20). This analysis is replaced in SEEGER's theory by the reasonable assumptions [B-8]

$$\begin{cases} W = H \\ \tau_0^{-1} = 2\pi\nu_0 . \end{cases} \tag{22}$$

In MASON's theory some of the above difficulties were avoided making use of rate theory; a similar treatment has more recently been given by SEEGER,

[9] See P. G. BORDONI, *Suppl. Nuovo Cimento*, **7**, 144 (1950).

DONTH and PFAFF [B-10] which do not assume the validity of the Arrhenius equation (18), but try to compute the energy exchanges between dislocations and lattice by means of the Kolmogoroff equation for diffusion processes.

The model for dislocation motion is the same that has been adopted by SEEGER, and the case of a suddenly applied stress is considered. As it has been already pointed out, when the energy of a pair of kinks exceeds a critical value, the two kinks are pulled apart by the applied stress. Owing to the statistical character of the energy exchanges between dislocations and lattice, the behaviour of the formers can be represented by the diffusion of their representative points in energy space. It may be observed that the density of such points must vanish when their energy exceeds the critical value for the applied stress. The average time required by the dislocation to reach this critical energy H_{cr} is the relaxation time. The estimation for the energy barrier H_{cr} is

$$H_{cr} = 2H_k\left[1-\frac{\pi\sigma}{8\sigma_p^0}\right] = \frac{4}{\pi}\sqrt{\frac{2}{5\pi}}\,\mu b^3\left(\frac{\sigma_p^0}{\mu}\right)^{\frac{1}{2}}\left[1-\frac{\pi\sigma}{8\sigma_p^0}\right], \tag{23}$$

which is near to one half of the value given by (17).

The relaxation time of equation (2) is then given by [10]

$$\tau = \frac{16}{\sqrt{5}\pi^3}\,\frac{b^4\mu}{c_t kT}\exp\left[-F_1\right], \tag{24}$$

where F_1 is a function of the two variables $(2H_k/kT)$ and $(H_{cr}/2H_k)$. The numerical values of F_1 are given by DONTH in a graph.

The maximum dissipation Q_m^{-1} is also related to the volume density N_0 of dislocations of lenght l by the approximate relation

$$Q_m^{-1} \simeq \frac{l^3 N_0}{24}. \tag{25}$$

From a qualitative standpoint it is quite remarkable that the temperature dependence of the relaxation time is not exponential, owing to the factor $(k\cdot T)^{-1}$ in equation (24) and to the relation between F_1 and $2H_k/kT$ which is not exactly linear. However for small stresses the ratio $H_{cr}/2H_k$ given by (23) is very near to unity. In this case the graph of the function F_1 is represented with good accuracy by

$$F_1 = 1-\frac{4H_k}{kT}. \tag{26}$$

---

[10] The coefficient of the exponential in eq. (24) is larger by a factor 2.4 than the coefficient $(2\pi B)^{-1}$ computed by means of the eq. (15) of the paper [A-16]. This difference is due to a different estimate for E_0 and c_t.

Moreover the variations of the function $(kT)^{-1}$ with temperature are small in comparison with those of the exponential term for all the temperatures of interest in the relaxation effect. If an intermediate temperature is chosen to compute the time factor in the right member of equation (24) (for instance $T = 100$ °K), the temperature dependence of relaxation time given by the theory of SEEGER, DONTH and PFAFF can be represented with a satisfactory approximation by the simple equation

$$\tau = \frac{16}{\sqrt{5}\pi^3 \cdot e \cdot 100 \cdot k} \frac{b^4 \mu}{c_t} \exp\left[\frac{4H_k}{kT}\right]. \tag{27}$$

The above equation shows that from a numerical standpoint this theory differs very little from the simpler one formerly given by SEEGER. As far as the activation energy is concerned, the difference between the value $4H_k$ taken in (27) and that $3.5\,H_k$ given by (17) is not physically significant, owing to the uncertainty in Peierls' stress. The same happens with the relaxation times

TABLE I. – *Physical properties of some f.c.c. metals involved in the relaxation theories.*

Metal	ϱ (g cm^{-3})	b (cm)	c_t (cm s^{-1})	μ (dyn cm^{-2})	σ_p^0/μ
Copper	8.96	$2.55 \cdot 10^{-8}$	$2.26 \cdot 10^5$	$4.6 \cdot 10^{11}$	$4.2 \cdot 10^{-4}$
Silver	10.49	2.88	1.6	2.9	6.0
Gold	19.3	2.88	1.2	2.85	8.6
Palladium	12.02	2.74	2.04	5.0	10.3
Platinum	21.45	2.76	1.7	6.1	3.2
Aluminium	2.70	2.85	3.08	2.64	5.0

computed for the two theories: of course in this case the comparison must be made between the logarithms, owing to the exponential character of equations (2) and (27). The values of the fundamental parameters required to compute τ_0^{-1} according to (20) and $(22)_2$ are listed in Table I. In the same

TABLE II. – *Comparison between the values of* $\ln \tau_0^{-1}$ *computed by means of the Seeger's theory and of the Seeger, Donth and Pfaff's theory.*

Metal	$\ln \tau_0^{-1}$		% difference
	Seeger theory	S.D.P. theory	
Copper	26.84	25.97	3.3
Silver	26.55	25.59	3.7
Gold	26.44	25.32	4.4
Palladium	27.12	25.49	6.3
Platinum	26.37	25.06	5.1
Aluminium	27.13	26.52	2.3

Table are also given the values of σ_p^0/μ computed by means of (17) and $(22)_1$ from the experimental value of W.

As is shown by Table II, the differences between the logarithms of τ_0^{-1} computed according to the Seeger's theory and to the S.D.P. theory are quite small. It may be added that the deviations of the exact relation (24) from the Arrhenius equation are too small to be detected with certainty even when the experiments are performed in a wide frequency range. From a numerical stand point the SEEGER's theory and the diffusion theory are equivalent, the main advantage of the latter residing in the smaller number of hypotheses required and in the possibility of evaluating the density of dislocations which give rise to the relaxation effect.

PART II. – Fundamental parameters of dislocation relaxation and influence of different treatments.

4. – Activation energy and characteristic frequency.

The validity of equation (4) may be tested plotting the logarithm of frequency against the inverse T_m^{-1} of the temperature of the peak: the experimental points must then lie on a straight line. The metals on which the dissipation peaks have been measured on specimens having the same purity and the same history, in a frequency range wide enough to make this control physically significant are aluminium, copper, silver, gold, palladium and platinum. For all these metals no systematic deviation from a straight line has been found, the slight scattering of experimental points being of the same order of the experimental accuracy (Figs. 22,

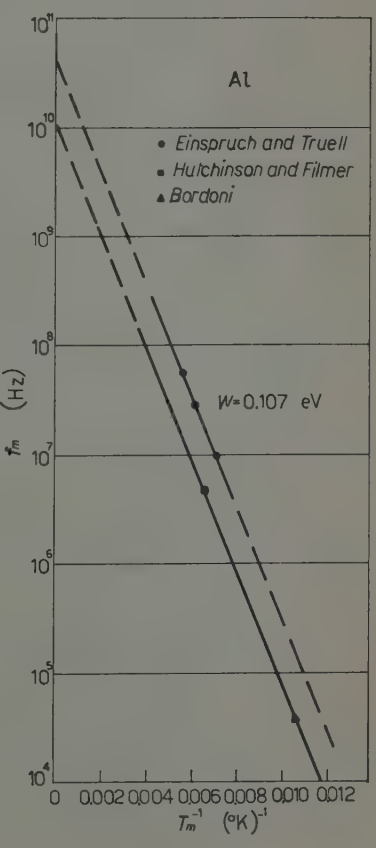


Fig. 22. – Frequency dependence of the temperature of maximum dissipation in polycrystalline aluminium: activation energy 0.107 eV [A-9].

23, 24). The data obtained for copper are particularly significant, owing to the wide frequency range covered by the measurements (from 1.8 kHz to 6.4 MHz) and to the fact that they were made not only on the same material, but actually on the same sample. Fig. 23 shows that the

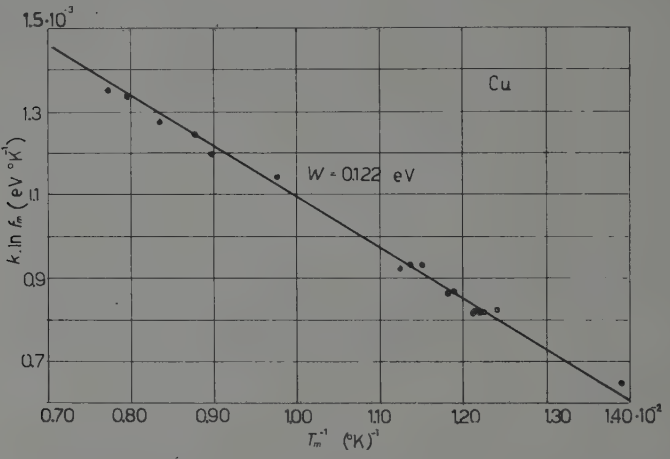


Fig. 23. – Frequency dependence of the temperature of maximum dissipation in technically pure polycrystalline copper: activation energy 0.122 eV [A-15, A-16].

parameters τ_0 and W have the same value for the flexural vibrations (lower frequencies) and for the longitudinal vibrations (higher frequencies), as the corresponding sets of experimental points lie on the *same* line. In fact for each type of vibrations the distance between the experimental points is large enough to show an eventual difference in the slope or position of the straight lines drawn separately through the points belonging to each type of vibration.

When the logarithm of frequency is multiplied by k, as it is done in

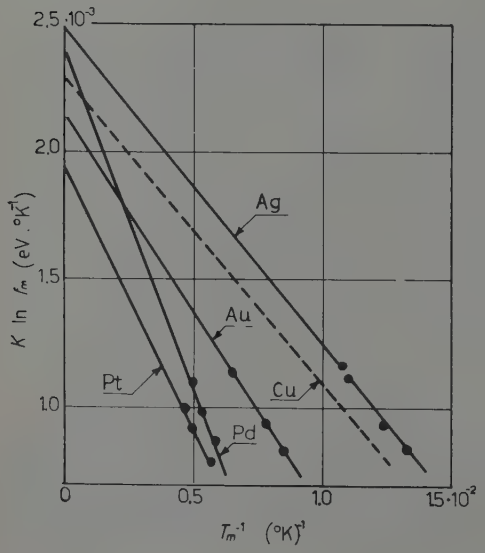


Fig. 24. – Frequency dependence of the temperature of maximum dissipation in silver, gold, palladium and platinum. Activation energies in eV: Ag, 0.124; Au, 0.158; Pd, 0.260; Pt, 0.192 [A-18].

Figs. 23, 24 the slope of the line gives the activation energy. The values obtained in this way are listed in Table III, together with the correspond-

TABLE III. – *Fundamental parameters of the relaxation effect.*

Metal	W (eV)	$\tau_0^{-1}/2\pi$ (s^{-1})	Debye frequency (s^{-1})	σ_p^0/μ	$\frac{2\pi c_t W}{\tau_0^{-1} b^4 \mu}$
Copper	0.122 ($\pm$0.005)	$3.8 \cdot 10^{11}$	$67 \cdot 10^{11}$	$4.2 \cdot 10^{-4}$	0.6
Silver	0.124 ($\pm$0.005)	40	45	6.0	0.04
Gold	0.158 ($\pm$0.002)	0.7	35	8.6	2.2
Palladium	0.260 ($\pm$0.013)	12.0	57	10.3	0.3
Platinum	0.192 ($\pm$0.006)	0.06	47	3.2	23
Aluminium	0.107 ($\pm$ —)	0.13	82	5.0	24

ing accuracy estimated from the scattering of experimental data. In the same Table are also given the values of $\tau_0^{-1}/2\pi$ computed from the intercept of the straight lines of Figs. 22, 23, 24, with the frequency axis, according to equation (2) and to the relation $2\pi f_m \tau = 1$. Hence $\tau_0^{-1} = \lim_{T \to \infty} 2\pi f_m$, and $\tau_0^{-1}/2\pi = f_{m,0}$. For comparison purposes the frequency associated with the Debye temperature ϑ has been computed for each metal from the equation $\nu = k\vartheta/h$, and is given in the next columm [11].

It is quite remarkable that the values taken by W in different metals are nearer to each other than could possibly be expected. This seems to support the idea of WEERTMAN and SEEGER, that the dislocation motion responsible for the effects, is associated with some intrinsic property of the dislocations, the differences between the potential barriers being small in crystals having

[11] The Debye frequencies associated with the different types of waves could be computed from the theory of specific heath of crystals, taking into account the Brillouin correction for wave lenghts of the same order as the lattice constant. However the values given in Table III are accurate enough to be compared with $\tau_0^{-1}/2\pi$.

the same structure. This result shows the interest of extending the investigation to crystals with a different structure.

The experimental value of $\tau_0^{-1}/2\pi$ is in every case smaller than the Debye's value. This is quite reasonable as the latter frequency is the highest that can be propagated in a periodic lattice. It must however be remembered that long before the melting point is reached the dislocations responsible for the relaxation effect are no longer active, as it is proved by the experiments on anneal (Figs. 11, 12, 13). To be physically significant the comparison ought to be made taking the value of the relaxation time not for an infinite temperature, but for the highest temperature that each metal can reach without cancelling the relaxation effect. To this purpose further experiments are needed on the influence of anneal on metals of known impurity content.

The ratio of the Peierls' stress to the shear modulus can be computed according to SEEGER's theory from equations (17), $(22)_1$ and from the experimental values of W. The values obtained in this way (Table III) are all higher than those obtained in static experiments. The explanation which has been given for copper, in which this difference was first noticed, can be applied to all the metals having the same structure. It has been observed [B-8] that a great portion of the dislocations dont lie along close-packed directions. Therefore their *effective* Peierls' stress which is observed in static experiments, is several orders of magnitude smaller than σ_p^0. This explanation removes the objection made by WEERTMAN [B-5], that the formation of kink pairs cannot be responsible for the relaxation effect, as the values of σ_p^0 computed by means of (10) or (17) are too large. It was suggested [A-11, A-13, B-8] to test SEEGER's theory using the values of σ_p^0/μ given by Table III to compute the dislocation frequency ν_0 by means of (20); this value could then be compared with that obtained experimentally for $\tau_0^{-1}/2\pi$. However the meaning of this comparison is not quite clear, as the preceeding theoretical discussion shows that the value of ν_0 given by (20) is probably of the same order as $\tau_0^{-1}/2\pi$, but does not necessarily coincide with it. A more significant test can be obtained eliminating σ_p^0/μ from (17) and (20), and observing that the expression $c_t H_s/\nu_0 b^4 \mu$ is an invariant for all metals. If we make the reasonable assumption that the unknown ratio between ν_0 and $\tau_0^{-1}/2\pi$ has the same value for all the metals having the same structure, the above expression must also be invariant when the experimental values of W and $\tau_0^{-1}/2\pi$ are substituted to H_s and ν_0. The last column of Table III shows that this is not the case, and that the evaluation of the relaxation time is actually the less satisfactory part of the theories of dislocation motion.

A further comparison has been made between the experimental data for copper and the theory of SEEGER, DONTH and PFAFF. A central point has been chosen on the line of Fig. 23, to compute the function F_1 and the ratio $2H_k/kT$ for the two cases $H_{cr}/2H_k = 1$ and $H_{cr}/2H_k = 0.9$. The frequency

766

of maximum attenuation has been evaluated as a function of the temperature by means of equation (24) and of the relation $2\pi f_m \tau = 1$. The comparison between the average experimental line and the theoretical points is shown in Fig. 25. The agreement may be considered satisfactory: however the theoretical points lie on a line somewhat steeper than the experimental one. A slight correction of the values taken by the function F_1 seems therefore to be required to obtain a better agreement between theory and experiment.

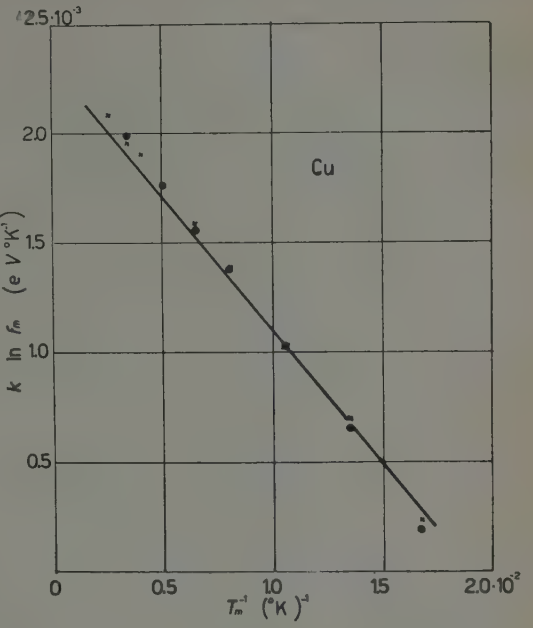


Fig. 25. – Comparison between the experimental data and the theory of Seeger, Donth and Pfaff for technically pure polycrystalline copper. Full circles, computed values for $H_{cr}/2H_k = 1$; crosses, computed values for $H_{cr}/2H_k = 0.9$. Heavy line, average of experimental data [A-16].

5. – Relaxation spectrum.

If the relaxation effect were due to the sum of many elementary effects with the same values for the activation energy and relaxation time, the experimental dissipation curves would coincide with those computed by Zener for a single relaxation time ([12]).

From equations (3) and (4) the ratio of Q^{-1} to its maximum value Q_m^{-1} can be expressed in a way independent of the frequency of measurement, as a function of the variable $W \cdot k^{-1} \cdot [T_m^{-1} - T^{-1}]$

$$\frac{Q^{-1}}{Q_m^{-1}} = \operatorname{sech} \frac{W}{k}\left[\frac{1}{T_m} - \frac{1}{T}\right], \tag{28}$$

where T_m is the temperature of maximum dissipation for a given frequency.

Taking for W the values given by Table III, the experimental ratio Q^{-1}/Q_m^{-1} has been compared with the curve (28). In every case the experi-

([12]) It must also be assumed that the ratio between the dislocation frequency and the parameter τ^{-1} which characterizes the relaxation of macroscopic anelastic strain is the same for all the dislocations. However this seems rather obvious from a physical standpoint.

mental lines are much broader than the theoretical curve for a single relaxation time, as it is shown by Figs. 26, 27 for the typical cases of copper and gold.

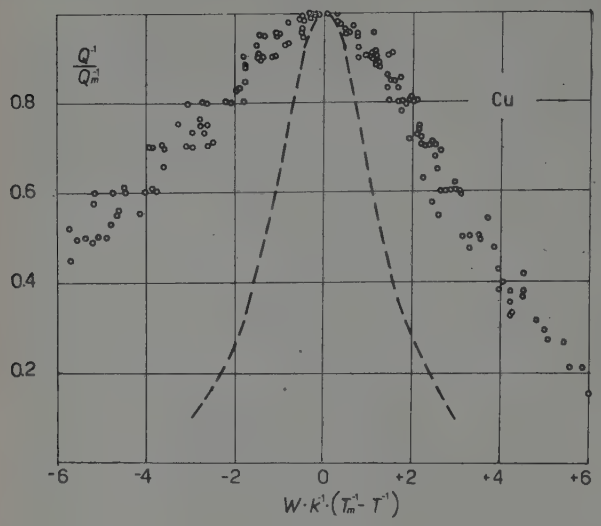


Fig. 26. – Values of Q^{-1}/Q_m^{-1} as a function of $W \cdot k^{-1} \cdot (T_m^{-1} - T^{-1})$ for polycrystalline copper. Open circles, experimental points; dotted line, Zener's dissipation curve for a single relaxation time [A-16].

Since the hypothesis of a single relaxation time must be discarded, the two simplest assumptions that can be made are:

a) all the elementary relaxation effects are associated with the same value of τ_0 but their activation energies are different;

b) all the elementary relaxation effects have the same activation energy W but the values for τ_0 are different.

Owing to the exponential relation (2) between τ and W, a constant logarithmic density of relaxation times between τ_1 and τ_2 corresponds to a constant linear density for energies between the values $(\overline{W} - \Delta W)$ and $(\overline{W} + \Delta W)$. As it has been shown [A-16], the value of Q^{-1} is given for an energy spec-

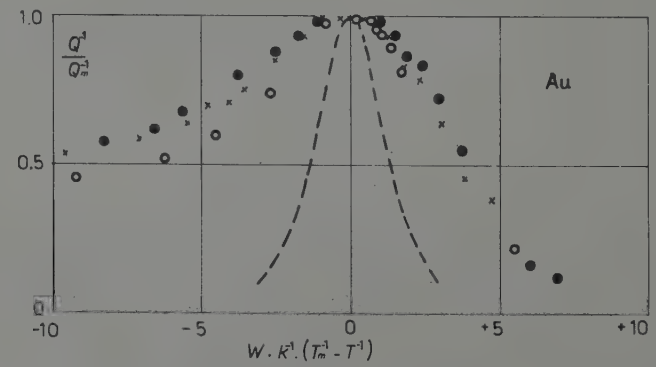


Fig. 27. – Values of Q^{-1}/Q_m^{-1} as a function of $W \cdot k^{-1} \cdot (T_m^{-1} - T^{-1})$ for polycrystalline gold. Full circles, open circles and crosses, experimental points; dotted line, Zener's dissipation curve for a single relaxation time [A-18].

trum with constant linear density by

$$Q^{-1} = \frac{S}{2\Delta W} kT \operatorname{tg}^{-1}\left[\overline{Y}_W \sinh \frac{\Delta W}{kT}\right], \tag{29}$$

where S is the total relaxation strenght $(M_U - M_R)/\sqrt{M_U \cdot M_R}$, M_U, M_R being the unrelaxed and relaxed values of elastic modulus; $2\Delta W$ is the width of the energy spectrum of center $\overline{W}$; $\overline{Y}_W$ is the attenuation function for a single relaxation time computed for the centre of the relaxation spectrum

$$\overline{Y}_W = \operatorname{sech}\left[\ln\left(2\pi f \overline{\tau}\right)\right] = \operatorname{sech}\left[\ln\left(2\pi f \tau_0 \exp\left[\overline{W}/kT\right]\right)\right]. \tag{30}$$

The dissipation given by (29) has been compared with the experimental results for copper, taking for $\overline{W}$ the value given by Table III and choosing

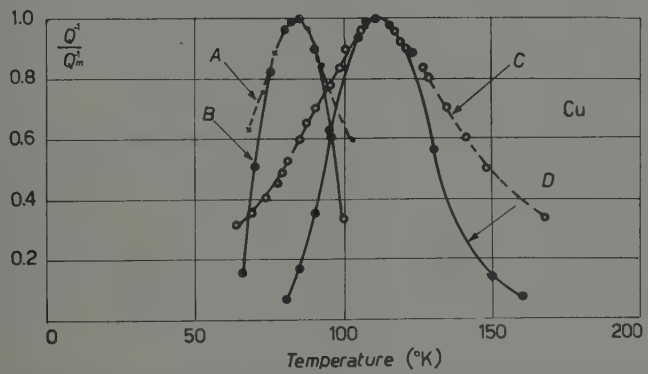


Fig. 28. – Comparison between experiment and theory for a relaxation spectrum with W variable and τ_0 constant. Curves A, C, measurements on polycrystalline copper at 13 kHz and 1 MHz. Curves B, D: computed values. $\overline{W}$ and τ_0 are given by Table III; $\Delta W = 0.019$ eV [A-16].

ΔW in order to make the theoretical curve as near as possible to the experimental dissipation values measured at about 13 kHz. The same values of $\overline{W}$ and ΔW have been employed to compute the dissipation for the same material at a frequency of 1 MHz. Fig. 28 shows that for the high frequency vibrations the computed curve is much narrower than the experimental one, and therefore the hypothesis a) does not seems to agree with the experimental data.

A more satisfactory result is obtained from the hypothesis b) of a constant logarithmic density of the values of τ_0. This hypothesis corresponds to a constant logarithmic density of times for every temperature, and the

dissipation is given by [A-16]

$$Q^{-1} = \frac{S}{2\eta} \operatorname{tg}^{-1} [\overline{Y}_\tau \sinh \eta] , \tag{31}$$

where S has the same meaning as in the previous equation; $\eta = \frac{1}{2} \ln \tau_2/\tau_1$. τ_2 and τ_1 being the upper and lower limits of the times spectrum, and $\overline{Y}_\tau$ is given by

$$\overline{Y}_\tau = \operatorname{sech} [\ln (2\pi f \overline{\tau})] = \operatorname{sech} [\ln (2\pi f \overline{\tau}_0 \exp [W/kT])] . \tag{32}$$

As it is shown by Fig. 29 it is possible to find a value of the parameter η characterizing the band width ($\eta = 2.70$) which makes the theorethical curves to agree fairly well with the experimental ones both at low and high frequency.

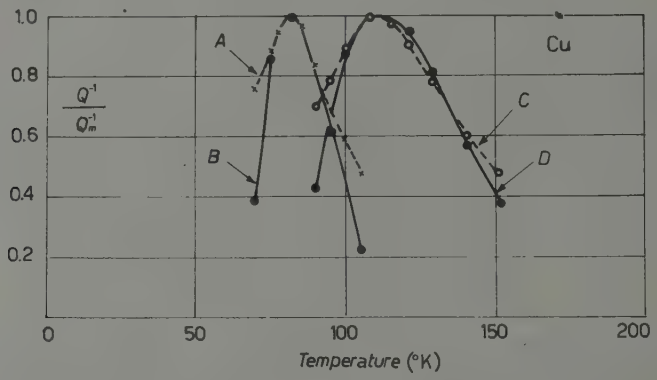


Fig. 29. - Comparison between experiment and theory for a relaxation spectrum with W constant and τ_0 variable. Curves A and C are the same as in Fig. 28; curves B, D have been computed for $\eta = 2.70$ [A-16].

To complete the comparison between the hypotheses on the different types of spectra it may observed that in the case *b*) the ratio Q^{-1}/Q_m^{-1} depends only upon the variable $Wk^{-1}[T_m^{-1} - T^{-1}]$, according to equation (31), as in the case of a single relaxation time. Hence when the experimental values of Q^{-1}/Q_m^{-1} are plotted as a function of the above variable, the points corresponding to all the vibration frequencies must lie on the same curve as it happens with the experimental points of the Figs. 26, 27. On the other hand this property is characteristic of the time spectra with constant energy. In fact equation (29) shows that for an energy spectrum the dissipation depends not only upon $\overline{W} \cdot k^{-1}[T_m^{-1} - T^{-1}]$ but also directly upon T and therefore upon the frequency of measurement. It may be concluded that the elementary relaxation effects responsible for the dissipation peaks differ in

the value of τ_0 but have the same value of the activation energy. To explain this result some modification is required in the present theories of dislocation motion, as τ_0 must obviously depend from a parameter which does not affect the value of the energy.

It is not surprising that the agreement between equation (31) and the experimental data becomes poorer when the distance from the peak increases. In fact the lower values of dissipation are considerably affected by the shape of the relaxation spectrum, which is only roughly represented by a constant logarithmic density. Moreover the effect of other causes of dissipation, different from dislocation relaxation, is more severely felt when the dissipation due to the latter cause is small.

It has been shown that no detailed information on the structure of the spectrum can be expected from measurements on relaxation effects, as their « resolving power » with respect to the spectral lines is comparatively poor ([13]). Some additional information about the shape of the relaxation spectrum can however be obtained, observing that the experimental points of Figs. 26, 27 are not symmetrical with respect to the peak, the points on the lower temperature side being somewhat higher than the corresponding points on the other side. As the differences are not large, it may be assumed that the main peak is associated with the symmetrical part of the curve fitting the experimental points. This symmetrical curve is quite well represented by the function

$$(33) \qquad \frac{Q^{-1}}{Q_m^{-1}} = \operatorname{sech}\left[\gamma \cdot \frac{W}{k}\left(\frac{1}{T_m} - \frac{1}{T}\right)\right],$$

where the parameter γ, which represents the spectrum width, is unity for

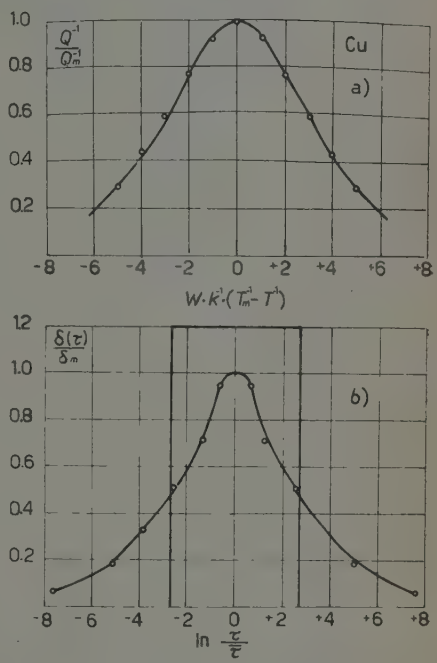


Fig. 30. – *a*) Comparison between the simmetrized values of dissipation in copper obtained from Fig. 26 (open circles) and the dissipation computed for a Fuoss-Kirkwood spectrum ($\gamma = 0.392$, heavy line). *b*) Fuoss-Kirkwood spectrum for $\gamma = 0.392$ and rectangular spectrum having the same total relaxation strenght and a bandwidth $\eta = 2.7$ [A-16].

([13]) See P. G. BORDONI: *Theory of relaxation effects with a continuous spectrum.* Proceedings of the III I.C.A. Congress, Stuttgart (September, 1959).

a single relaxation time, and vanishes for an infinite spectrum. According to the theory given by Fuoss and Kirkwood [14] the relaxation spectrum density $\delta(\tau)$ which gives rise to the dissipation (33) is given by

$$\delta(\tau) = \frac{\gamma S}{\pi} \frac{\cos(\gamma\pi/2)\cosh(\gamma \ln \tau/\bar{\tau})}{\cos^2(\gamma\pi/2) + \sinh^2(\gamma \ln \tau/\bar{\tau})}. \tag{34}$$

As it is shown by Fig. 30*a* the agreement between the theoretical line computed from (33) with $\gamma = 0.39$ and the experimental points obtained from the symmetrization of Fig. 26 is good, even for values of Q^{-1}/Q_m^{-1} as low as 0.3. The same computation has been made for the other metals whose dissipation measurements were accurate enough, and the values found for the parameter γ are collected in Table IV. The relaxation times τ_1, τ_2 for which the height of the spectrum is $1/\sqrt{2}$ of the maximum height have been computed in every case, and their ratios are listed in Table IV. When the values of τ_2/τ_1 are related to the treatments undergone by the metals it is found that the wider spectra correspond to a larger amount of cold work.

Table IV. – *Relaxation spectra.*

Metal	Silver	Gold	Palladium	Platinum	Copper
γ	0.266	0.345	0.284	0.426	0.390
τ_2/τ_1	335	57	205	18	28

6. – Frequency relaxation.

The experimental data concerning the temperature dependence of the resonant frequency for copper [A-16] are accurate enough to allow a comparison between the measured values of $(f - f_0)/f_0$ and those computed according to the relaxation theory. As in equation (6), f is the frequency at the temperature T, while f_0 is the relaxed value of the resonant frequency.

The above ratio can be obtained from the experimental data subtracting from the frequency measured at the temperature T, the frequency computed with a linear extrapolation from room temperature measurements. In this way the monotonic frequency changes due to the effect of thermal expansion are cancelled and only the effects of relaxation are left [15].

[14] See R. M. Fuoss and J. G. Kirkwood: *Journ. Chem. Phys.*, **63**, 385 (1941).
[15] For the slight errors introduced by the linear extrapolation see [A-16].

The equation (6) is no longer valid, the effect being due to a relaxation spectrum of considerable width, as it has been shown by the dissipation measurements. For a constant logarithmic density of times, the ratio $(f-f_0)/f_0$ is given by

$$\frac{f-f_0}{f_0} = \frac{S}{2}\left\{\frac{1}{2} + \frac{1}{4\eta}\ln\frac{\cosh\left[\eta - Wk^{-1}\left(T_m^{-1} - T^{-1}\right)\right]}{\cosh\left[\eta + Wk^{-1}\left(T_m^{-1} - T^{-1}\right)\right]}\right\}. \tag{35}$$

It must be noticed that $(f-f_0)/f_0$ for vanishing T is independent of the width η of the spectrum, and equals $S/2$ as in the case of a single relaxation time. The same does not happen for the maximum dissipation Q_m^{-1} which is a function of the spectrum width, and is given by

$$Q_m^{-1} = \frac{S}{2}\cdot\frac{\operatorname{tg}^{-1}(\sinh\eta)}{\eta} \leqslant \frac{S}{2}, \tag{36}$$

according to (31).

The equivalent of equation (35) cannot be obtained in closed form for a Fuoss-Kirkwood spectrum; however it is easily proved that the limiting value of $(f-f_0)/f_0$ when T vanishes is still $S/2$, whilst the maximum dissipation depends upon the spectrum width as in the former case, and is given by

$$Q_m^{-1} = \gamma\frac{S}{2} \leqslant \frac{S}{2}. \tag{37}$$

The experimental values of $(f-f_0)/f_0$ are plotted in Fig. 31 for three different types of vibrations. The corresponding dissipation curves are also plotted in the same scale for comparison purposes. Two facts are immediately observed: *a*) the total frequency relaxation is much smaller for

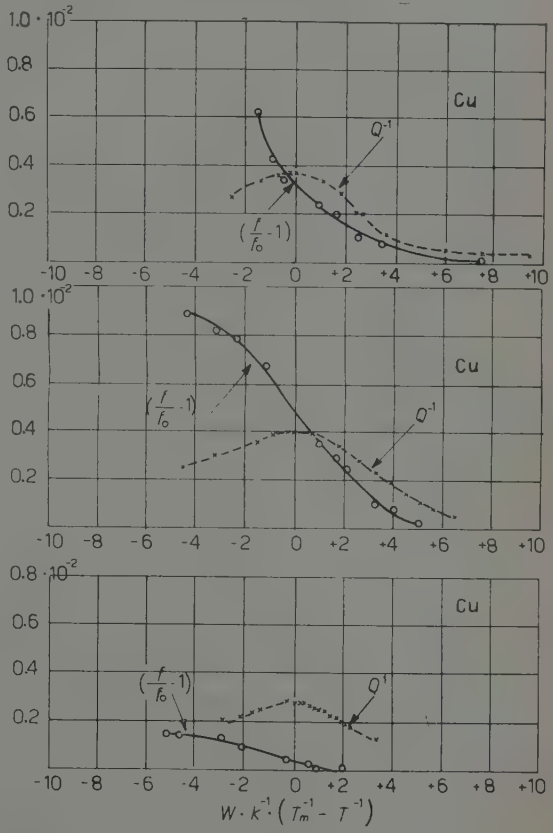


Fig. 31. - Frequency-relaxation and dissipation for different vibration frequencies. Abscissae $W\cdot k^{-1}\cdot(T_m^{-1}-T^{-1})$; ordinates Q^{-1} and $(f-f_0)/f_0$. Upper curves, 1.8 kHz; middle curves, 13 kHz; lower curves, 1 MHz. Polycrystalline copper, technically pure [A-16].

longitudinal vibrations (lower diagram) then for the flexural vibrations of plates and rods (intermediate and upper diagram); *b*) for longitudinal vibrations the total frequency relaxation is smaller than the corresponding Q_m^{-1}, in contrast with equations (36) and (37).

It is not difficult to explain *a*) observing that the dislocation motion changes the shape of the solid rather than its volume. As it was suggested by ZENER it may be assumed that the compressibility modulus is not affected by the relaxation effect. The classical relations between the elastic moduli give then rise to the following equation between the relaxation strenght S_l for longitudinal vibrations and the corresponding value S_e for flexural or extensional vibrations

$$S_l = 2S_e \frac{c_t^4}{c_e^2 c_l^2}, \tag{38}$$

where c_t, c_e, c_l are the velocities for torsional, extensional and longitudinal waves. Taking for them the values given by the literature, the value 0.19 is found for the ratio S_l/S_e in good agreement with the experimental curves of Fig. 31. This result shows the interest of a more extended experimental investigation of this point, as the property of not affecting the compression modulus seems to be quite important for the mechanism of dislocation motion.

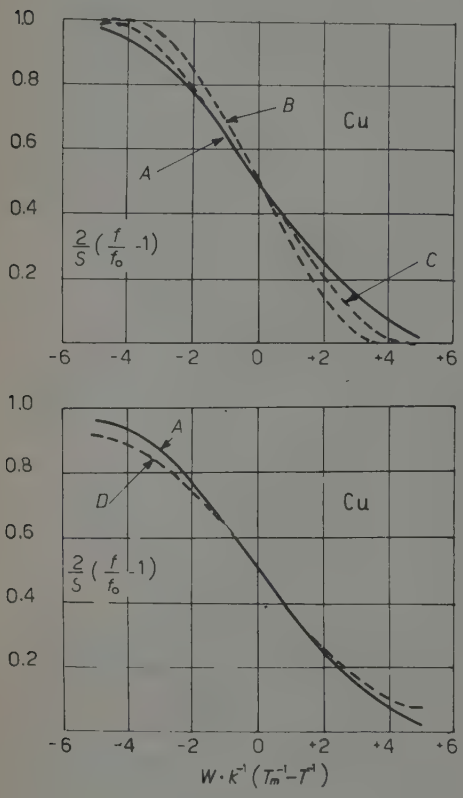


Fig. 32. – Comparison between the experimental relaxation of frequency, in technically pure copper, at 13 kHz (Curve *A*) and the theoretical curves for a constant logarithmic density (*B*, $\eta=2.70$; *C*, $\eta=3.5$) or for a Fuoss-Kirkwood spectrum (*D*, $\gamma = 0.392$) [A-16].

Point *b*) is somewhat less easy to explain as it is in open contradiction with theory; it must however be remembered that the measurements were made on polycrystalline copper. The stress due to the high frequency vibrations may be considered as the sum of a purely longitudinal stress and of a more complex stress changing from grain to grain, owing to the large elastic anysotropy of copper crystals and to their random orientation. The latter stress does

not correspond to any appreciable macroscopic force, because of the randomness of grain orientation, and therefore has no influence upon the vibration frequency. On the contrary the elastic energy associated with the same stress is not negligible in comparison with the energy associated with the longitudinal part, and must be taken into account when the total dissipation is evaluated. Hence the dissipation is larger than the value corresponding to a purely longitudinal stress, as it is found in the lower diagram of Fig. 31. The comparison between the experimental values of $(f - f_0)/f_0$ and those computed for a constant logarithmic density or for a Fuoss-Kirkwood spectrum (Fig. 32) shows that the latter type of spectrum gives a better approximation, as it is found in dissipation measurements.

7. – Subsidiary dissipation peak.

A second peak was found by NIBLETT and WILKS below the temperature of the main dissipation maximum, in high purity (99.999%) polycrystalline copper, at vibration frequencies between 0.38 and 1.1 kHz, as it is shown by Figs. 12, 17 [A-4, A-8]. From the original measurements the separation of the second peak from the main one does not seem to depend upon mechanical or thermal treatments, neutron irradiation or strain amplitude. The same peak is also clearly shown at somewhat higher temperatures by PARÈ's measurements on copper single crystals of the same purity (99.999%) for a frequency range between 3.7 and 5.6 kHz [A-13]. In measurements made at higher frequencies the N.W. peak is less clearly exhibited, being rather a flat « bump » on the low temperature side of the curve (Fig. 8) even when the measurements are made on high purity specimens (Fig. 19). In the last measurements the separation of the two peaks is smaller than in the low frequency experiments.

The above results show that the N.W. peak is due to a thermally activated relaxation effect, with an activation energy W' smaller than the energy W of the main peak. To evaluate W', the N.W. peak must be separated from the main maximum. This can be done computing the $(\cosh^{-1})$ of the experimental values of Q_m^{-1}/Q^{-1} and plotting it as a function of $W \cdot k^{-1}[T_m^{-1} - T^{-1}]$. According to the previous remarks on the shape of the relaxation spectrum and on the Fuoss-Kirkwood approximation, the curves obtained in this way are straight lines near the main maximum and on the high temperature side, where the effects of the secondary maximum are little felt. The N.W. peak can be isolated taking the difference between the straight line, obtained by extrapolation of the values near the main maximum, and the low-temperature branch of the experimental curve. The hyperbolic secant of these

values multiplied by Q_m^{-1} gives the dissipation due to the N.W. peak. Fig. 33 clearly shows the increase of the temperature T'_m of the subsidiary peak with the frequency f_m. Similar computations have been made for other measurements, and the frequencies have been plotted in a logarithmic scale against $(T'_m)^{-1}$ in Fig. 34. The experimental points are placed on a straight (solid) line, and their agreement with the values obtained by other experimenters is reasonably good [A-13]. The activation energy computed from the slope of the line has the value $W' = 0.041$ eV being about $\frac{1}{3}$ of the energy found for the main peak in the same metal (Table III). The value of $(\tau'_0)^{-1}$ computed from the intercept of the experimental line with the frequency axis is smaller for the N.W. peak than for the main peak. As it is shown by Fig. 34, the vibration frequencies corresponding to the two peaks coincide at room temperature.

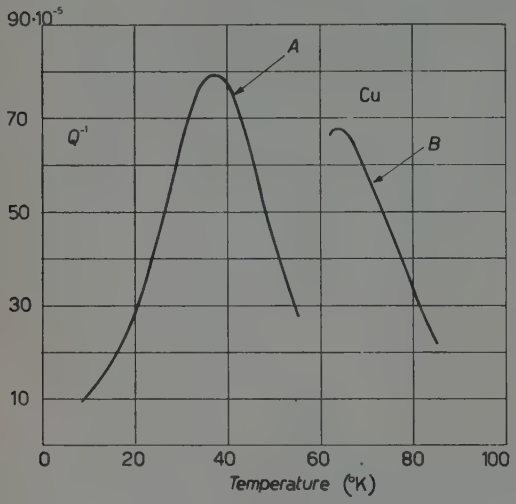


Fig. 33. – Niblett and Wilks peak. Polycrystalline copper technically pure. Curve *A*, 15 kHz; *B*, 1.91 MHz [A-16].

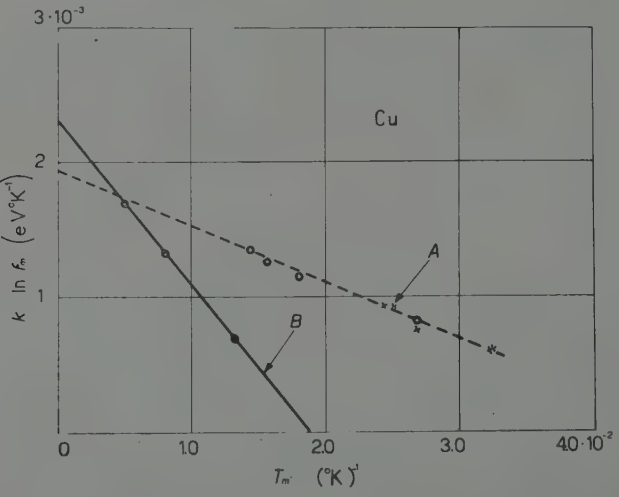


Fig. 34. – Frequency dependence of the temperature of the N. W. peak (line *A*) and of the main peak (line *B*) in polycrystalline copper. Open circles [A-16]; crosses, (A-8, A-13). $W' = 0.041$ eV [A-16].

The comparatively low value found for W' supports SEEGER's explanation of the second peak [B-8]. In f.c.c. metals the dislocations lines which are responsible for the low temperature attenuation lie in the {111} glide planes along one of the close-packed directions, and have Burger's vectors $\frac{1}{2}\langle 110\rangle$. The angle between the dislocation line and the Burger's vector may then be either $\pm 60°$, or $0°$. Theory shows that the energy for the formation of a kink pair must be different for the two cases, and according to (17) and (22) this gives rise to different values for the activation energy of the relaxation effect.

It may be added that recent investigations in silver, gold, palladium and platinum [A-18, A-20] have shown that the low-temperature side of the attenuation curves is always higher than the other side when the measured values are plotted as a function of $W \cdot k^{-1}[T_m^{-1} - T^{-1}]$, as it is shown by Fig. 27 in the case of gold.

This result seems to indicate the presence of a subsidiary dissipation maximum near to the main one. The N.W. peak is therefore a rather general, if not very evident, feature of the dissipation-temperature curves for f.c.c. metals. Further experiments at lower frequencies may consent the separation of the two peaks, as their activation energies are probably different, as it has been found in copper.

8. – Influence of anneal, cold-work, impurities, strain amplitude and neutron irradiation.

Experiment shows that the non-linear behaviour of metals is closely related to the thermal and mechanical treatments. On the other hand the effects of these treatments and of neutron irradiation are strongly dependent upon the impurity content. It is therefore convenient to consider together the influence on the relaxation effect of all the above causes, notwithstanding their apparent eterogeneity.

It has already been shown that an anneal at moderate temperatures reduces the height of the relaxation peak: in addition to this effect an increase of the room temperature dissipation is produced when the temperature of the treatment is raised. The second effect is generally not found in impure metals, as it was the case with some of the copper and silver specimens on which the first measurements were made [A-1]. In this case the peak can be easily cancelled by anneal, obtaining a low value of dissipation even at room temperature, as it is shown by Fig. 11 in the case of commercial copper after a 2 hours treatment at 175 °C. A similar treatment (1 hour at 180 °C) seems to be less effective in cancelling the peak in high purity (99.999%) copper (Fig. 12). If the temperature of the treatment is raised, the peak

disappears but a considerable increase is noticed in the dissipation at room temperature, as it is also shown for a somewhat less pure copper (99.965%) by curve *A* of Fig. 17. All the above experiments were made on polycrystalline specimens; however CASWELL [A-11] and THOMPSON and HOLMES [A-17]

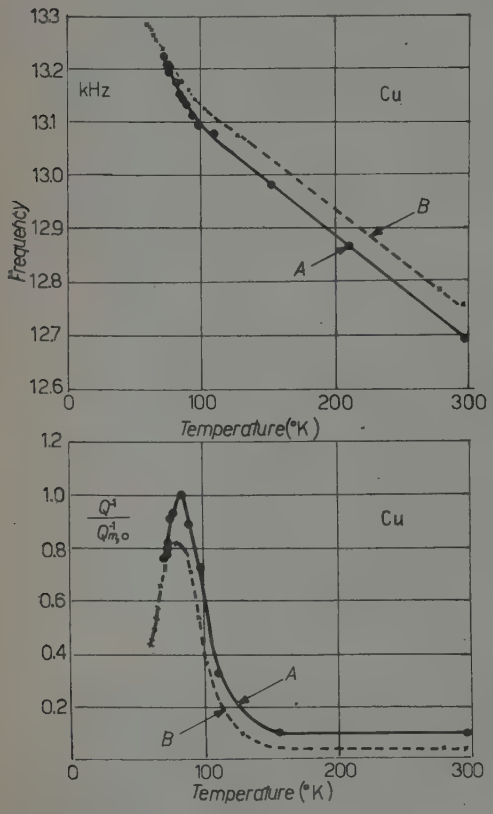


Fig. 35. – Influence of anneal on polycrystalline technically pure copper. *A*, after machining; *B*, after 16 hours at 138 °C [A-16].

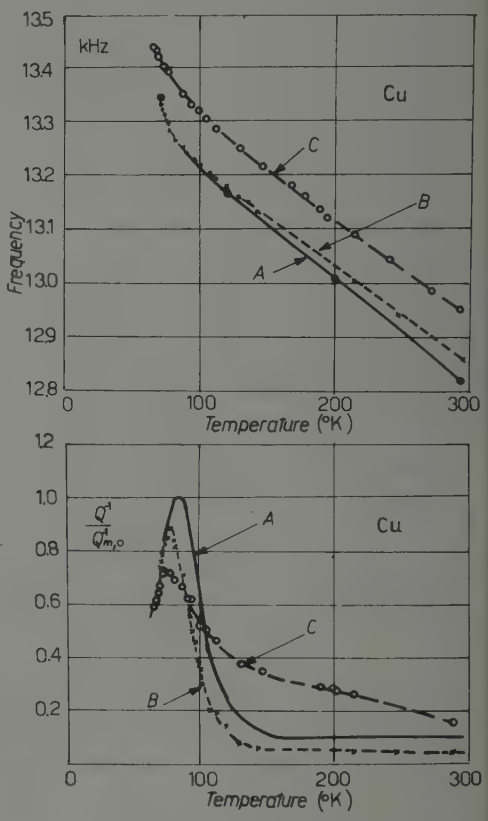


Fig. 36. – Influence of anneal on polycrystalline technically pure copper. *A*, after machining; *B*, after 2 hours at 225 °C; *C*, after an additional treatment for 1 hour at 242 °C [A-16].

have found that both effects of anneal are present in single crystals, and the same happens with polycrystalline 99.80% pure silver, as it shown by Fig. 13.

To get further information about the two effects of anneal, the dependence of the resonant frequency and of Q^{-1} upon temperature has been systematically investigated for polycrystalline 99.80% pure copper after different thermal treatments [A-16]. The reduction of the height of the peak is the only

effect produced when the temperature of anneal does not exceed 225 °C (Figs. 35, 36, curves *A*, *B*). The same treatment reduces also the inflection of the frequency-temperature curve, without changing $\bar{\tau}_0$. The above changes in dissipation and frequency are easily explained assuming that the number of

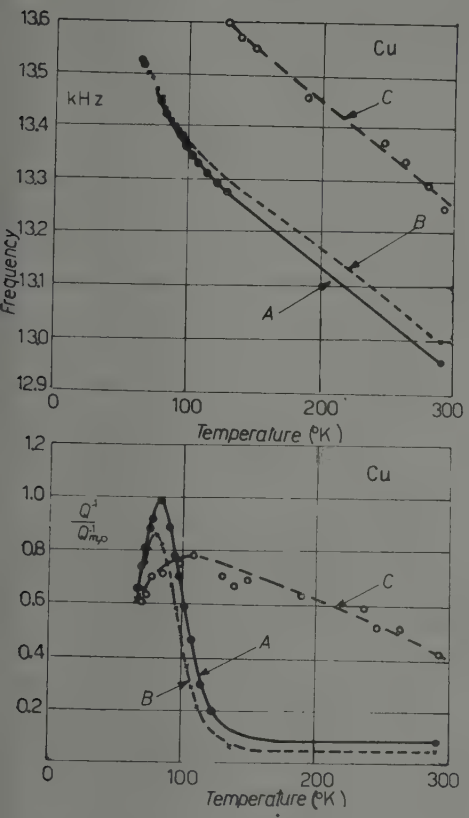


Fig. 37. – Influence of anneal on polycrystalline technically pure copper. *A*, after machining; *B*, after 5 hours at 146 °C; *C*, after an additional treatment for 6 hours at 243 °C [A-16].

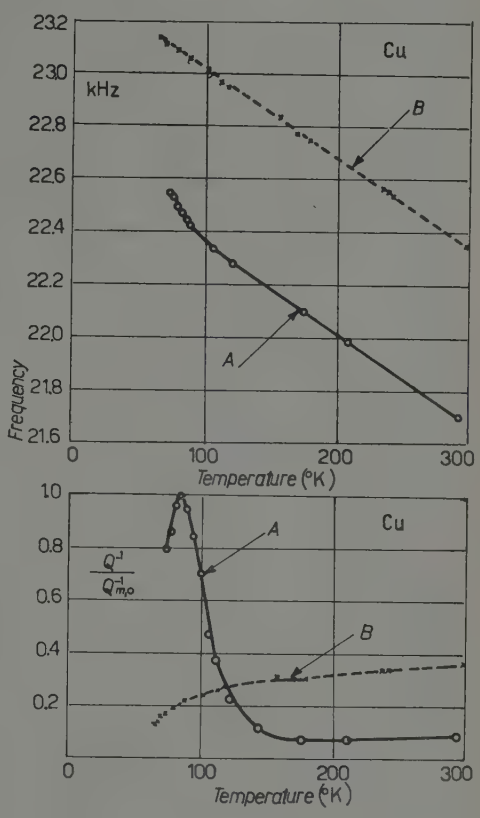


Fig. 38. – Influence of anneal on polycrystalline technically pure copper. *A*, after machining; *B*, after 1 hour at 600 °C [A-16].

dislocations active in the relaxation effect is decreased by a low temperature anneal.

When the annealing temperature exceeds 225 °C, a more drastic reduction of the peak is obtained, but the shape of the dissipation curve is entirely changed, and the room temperature dissipation is increased; in the same time the frequency is raised by an amount independent of temperature (Fig. 36, 37, curves *C*; Fig. 38, curve *B*). After an high temperature anneal (1 hour

at 600 °C) the frequency and the dissipation are very sensitive to the strain amplitude, even for strains of the order of 10^{-7} (Fig. 39). The same amplitude dependence produced by an high temperature anneal has also been found in copper single crystals [A-11] and in polycrystalline silver [A-20]. It may be observed that the curves C in the dissipation diagrams of Figs. 36, 37 and the curve B in the same diagram of Fig. 38, were obtained for strains of the order of 10^{-8} and can therefore be considered as the limiting values of the dissipation for vanishing strain.

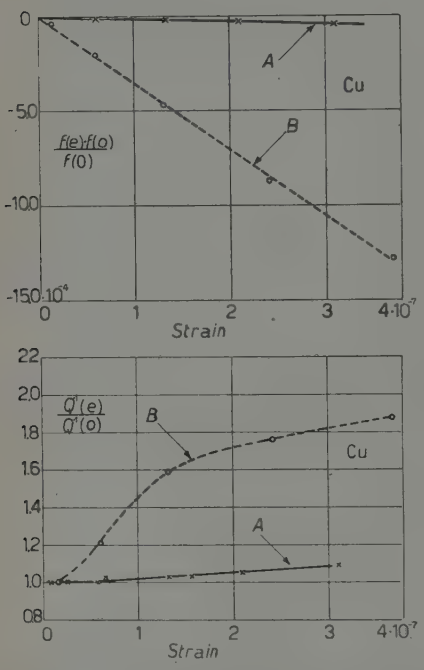


Fig. 39. – Strain amplitude dependence of frequency and dissipation at room temperature, in polycrystalline technically pure copper. A, after machining; B, after 1 hour at 600 °C [A-16].

At the present time no theory exists which adequately describes all the effects of an high temperature anneal. From the temperature of the treatment it could be suspected that some of these effects are due to recrystallization, but it must be remembered that the high value of room temperature dissipation has been found also in single crystals. As it has been pointed out [A-11], the dissipation measured after a high temperature anneal seems to be independent of frequency, at least in the range covered by the available data ((0.5÷40) kHz)). It would certainly be interesting to know if the temperature-independent increase of frequency shown by Figs. 36, 37, 38 which makes the explanation of the effect more difficult, is also observed in single crystals.

All the available experimental data show that the peak is not only reduced by a low-temperature anneal but also slightly shifted towards the lower temperatures, as it may be seen from Figs. 11, 12, 13, 35, 36, 37. This indicates that the relaxation effects more easily cancelled are those due to the dislocations associated with higher values of relaxation times. In fact, after the treatment, the centre of the relaxation times spectrum is lower than before, as it is shown by the slight decrease of T_m. Measurements on copper have shown that $\bar{\tau}_0$ may be reduced by a factor of 0.6; larger changes in the times spectrum can be expected in silver, in which the temperature shift of the peak is very evident (Fig. 13) according to recent experiments [A-20].

780

The effect of different amounts of cold work upon the height of the main and subsidiary peak has been investigated by NIBLETT and WILKS in 99.999% pure polycrystalline copper [A-8]. When the prestrain is of the order of 0.1% the peaks are hardly noticed; their height increases with the amount of the prestrain, whilst the room temperature attenuation is reduced (Fig. 40). When the prestrain exceeds the value 2%, the amplitude of both peaks becomes approximately independent of the amount of prestrain. A similar investigation has been made by CASWELL on a 99.999% pure copper single crystal [A-11]. The results agree with the preceding ones, the amount of permanent strain required to reach the maximum height of the peak being somewhat higher in the latter case, as it is shown by Fig. 19. Moreover CASWELL's experiments show that cross-rolling has a more pronounced effect on the height of the peak than continued rolling in one direction. This may be possibly due to the fact that cross-rolling activates different sets of slip planes.

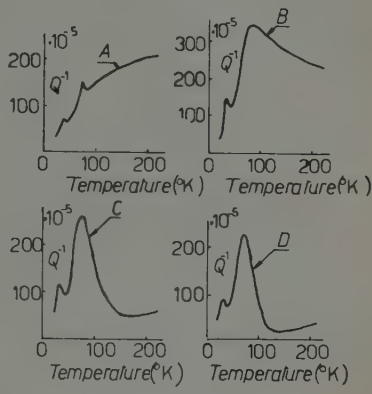


Fig. 40. – Dependence of Q_m^{-1} upon cold work for a 99.999% pure polycrystalline copper [A-8]; curve *A*, strained 0.1%; *B*, 0.5%; *C*, 2.2%; *D*, 8.4%.

The dependence of maximum dissipation upon prestrain, can be explained qualitatively [B-8] observing that the dislocations active in the relaxation effect are approximately parallel to a densely packed direction. This is only a small fraction of the total number of dislocations; as their number increases with increasing prestrain, the active fraction generally decreases, owing to the elastic forces between dislocations. These forces increase with the number of dislocations and push them into directions different from the close packed ones.

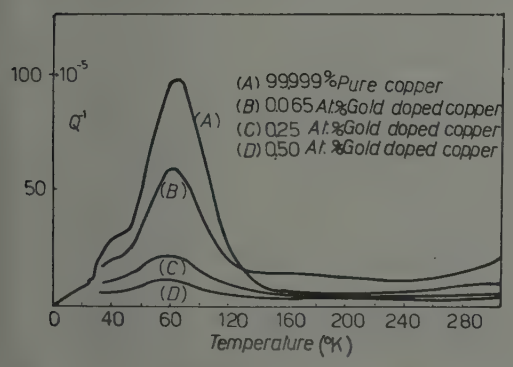


Fig. 41. – Dependence of Q_m^{-1} on gold atomic %, in a cross-rolled copper crystal [A-11].

For a small amount of cold work the increase in the total number of dislocations predominates, and the height of the peak increases. However the total number of dislocations parallel to close-packed directions soon reaches a saturation value under the controlling effect of the

elastic interaction, and the peak becomes independent of the amount of prestrain.

The effects of cross-rolling have also been compared for several copper crystals containing a different amount of gold [A-11]. As it shown by Fig. 41 the height of the peak is very sensitive to the presence of gold atoms, being

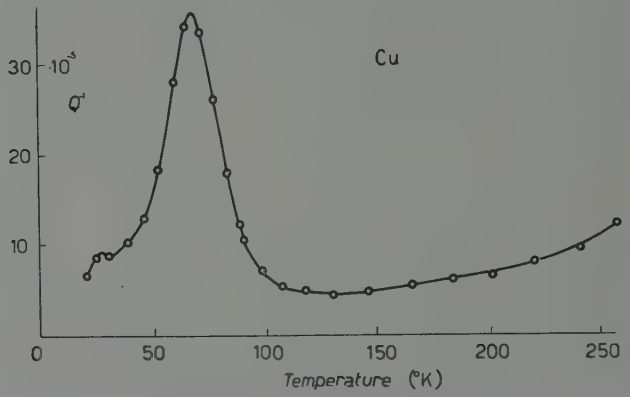


Fig. 42. – Dissipation in polycrystalline copper 99.999% pure, strained 6.6%, and neutron irradiated ($5 \cdot 10^{18}$ n.v.t.) [A-8].

reduced to about $\frac{1}{2}$ of its value in the 99.999% pure crystal, by 0,065% atomic % of gold, and being almost cancelled by 0.5 atomic % gold. The original measurements [A-11] show that this reduction of the peak takes place without any change in its temperature, in complete agreement with the fundamental hypotheses of dislocation theory, which have been discussed above. The number of foreign atoms seems to control the number of dislocations active in the relaxation effect, but has little or no influence on the activation energy and on the time τ_0.

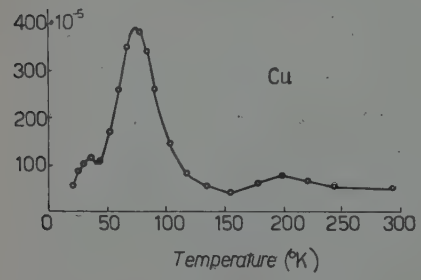


Fig. 43. – Dissipation in polycrystalline copper 99.999% pure, neutron irradiated ($5 \cdot 10^{18}$ n.v.t.), and subsequently strained 5.1% [A-8].

The influence of a fast neutron bombardament on dissipation has been investigated on a 99.999% pure polycrystalline copper for a radiation flux of $5 \cdot 10^{18}$ n. v. t. [A-8]. A substantial reduction of the height of the peak is found, together with a slight shift towards the lower temperatures, as after an anneal. This may be seen comparing the peak height of Fig. 42, which is of the order of $35 \cdot 10^{-5}$, with the peak for the same specimen (poly-

crystalline copper 99.999% pure) before the irradiation, whose height, given by Fig. 12, is about $240 \cdot 10^{-5}$. When the sample is strained after irradiation, the height of the peak may be larger than for the unirradiated material (Fig. 43). Other measurements made on a polycrystalline specimen of commercial copper (purity not stated) for a smaller irradiation flux ($5 \cdot 10^{17}$ n. v. t.) are in qualitative agreement with those of Fig. 42 [A-17]. However the peak reduction is only of 20%, and no temperature shift is found when the two curves are superposed, as it may be seen from Fig. 44. The neutron bombardment seems to have a larger effect on high purity copper crystals which have previously undergone a complete anneal at 650 °C. The high attenuation found at room temperature after anneal is reduced to 1/10 of its value by an irradiation flux on only $2 \cdot 10^{12}$ n. v. t. while a small relaxation peak is faintly shown (Fig. 45). At the

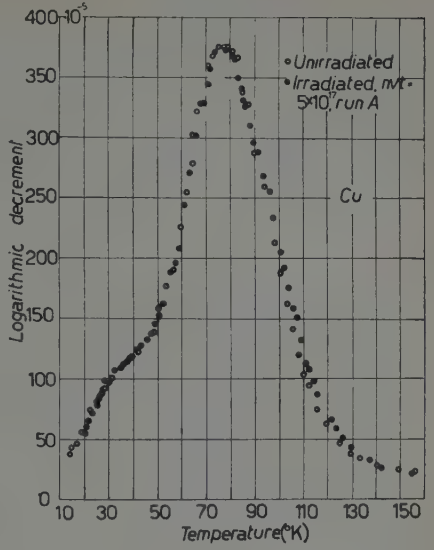


Fig. 44. – Dissipation in technically pure polycrystalline copper. The unirradiated values have been reduced by a factor of about 20% [A-17].

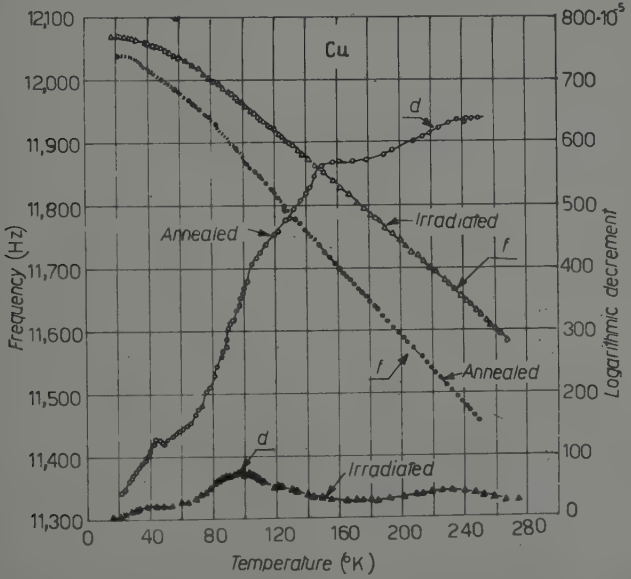


Fig. 45. – Influence of neutron irradiation upon frequency and dissipation of a 99.999% pure copper crystal, previously annealed [A-17].

same time the room temperature frequency is increased by an amount of about 2%. This result supports the hypothesis that high energy irradiation can reduce the number of active dislocations, as it is done by anneal and by increasing impurity content, but does not affect the basic parameters W and W'. Moreover the influence of irradiation seems to be strongly dependent upon the presence of foreign atoms, being rather small for impure metals.

9. – Relaxation peak in alloys.

The experimental investigation of the relaxation effect has been recently extended to a complete system of gold-silver alloys, which has been chosen owing to the complete solubility in the solid phase, due to the favourable ratio of the atomic diameters (1.0014 for silver-gold).

To make the results more easily comparable, all the specimens have undergone the same thermal and mechanical treatments, and their size has been chosen in such a way the all their resonant frequencies are near to 20 kHz. As it is shown by Fig. 46 the peak is clearly exhibited by the systems with a small concentration of foreign atoms, and reduces to a flat bump when the atomic concentration approaches 50%. This may be clearly seen in Fig. 47, where the height Q_m^{-1} of the peak is plotted as a function of the atomic percent of gold.

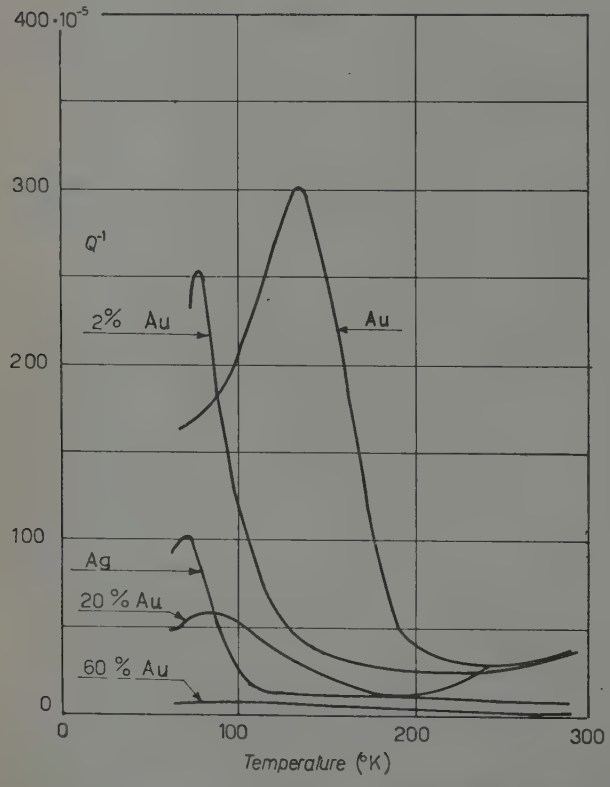


Fig. 46. – Dissipation in silver-gold alloys: resonant frequency $\simeq$ 20 kHz [A-18].

On the silver side the values of Q_m^{-1} have a maximum for a concentration of about 2 atomic % of gold, the height of the peak for this alloy being more than twice that for pure silver. This result shows that a small

concentration of impurities may favor the formation of active dislocations by cold work. This explanation agrees with the fact that the room temperature velocity of extensional waves c_e has a minimum for the same alloy (Fig. 48), owing to the larger number of dislocations which can move freely at that temperature. It must also be noticed that the temperature coefficient of frequency shows an anomalous behaviour for the same concentration.

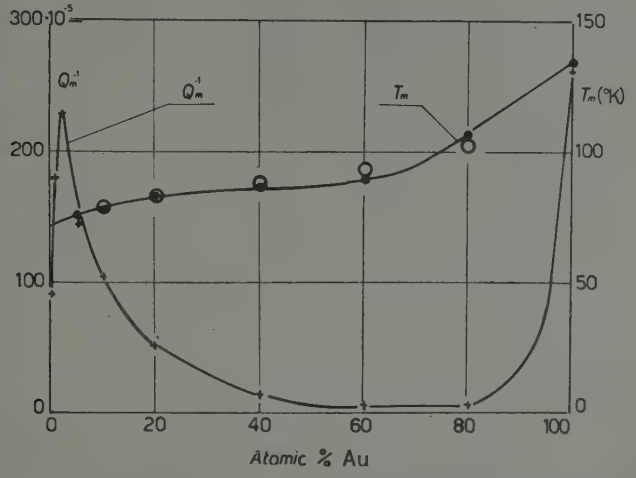


Fig. 47. – Values of Q_m^{-1} and T_m as a function of gold atomic %, in silver-gold alloys. Open circles, computed values of T_m [A-18].

For small concentrations of gold it is quite interesting to compare the Q_m^{-1}-concentration curve of Fig. 47, with the similar curve of Fig. 41, for a small concentration of gold in copper crystals. It is not surprising that a maximum for Q_m^{-1} is not found in the second curve. In fact owing to the less favorable ratio of the atomic diameters (1.1283 for gold-copper), corresponding to a stronger lattice distortion for the same atomic concentration, the maximum for gold-copper alloys can eventually be expected for gold concentrations smaller than 0.065 atomic percent, and therefore may have escaped the notice of the experimenters. On the other hand, it is also possible that the maximum in the Q_m^{-1}-concentration curve may only be found for atomic ratios very near to unity. For this reason it would be interesting to extend the investigation to small concentrations of silver-gold on the gold side, between the last alloy considered in Fig. 45 (80 atomic % gold) and the pure metal.

The temperature T_m of the peak for a vibration frequency of about 20 kHz rises gradually with the gold concentration from the value 71.5 °K for pure silver to 135 °K for pure gold. This is an additional proof of the intrinsic character of the activation energy and of $\bar{\tau}_0$ which are very little influenced

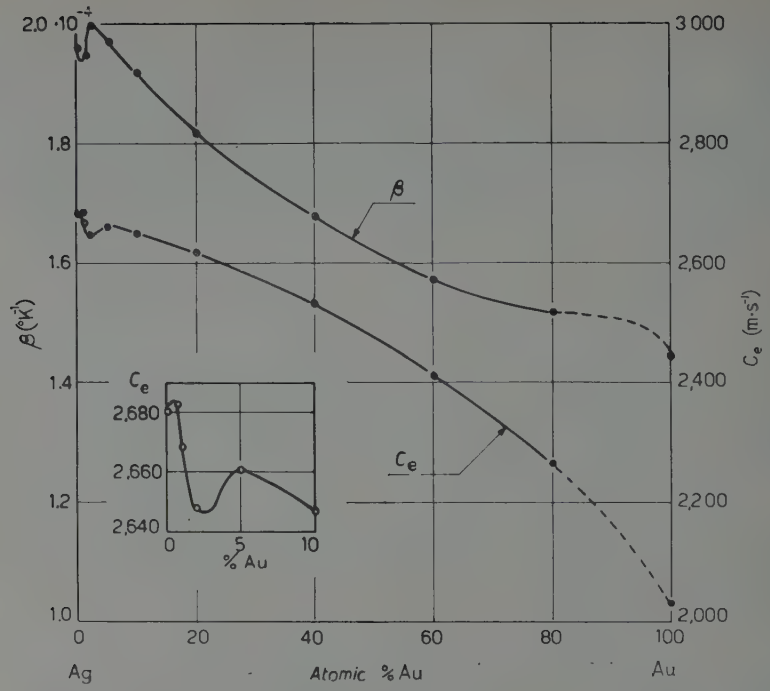


Fig. 48. – Values of the extensional velocity c_e and of the frequency-temperature coefficient $\beta = -\partial \ln f/\partial T$ measured at room temperature, as a function of gold atomic %, in silver-gold alloys [A-18].

by the presence of foreign atoms. These atoms seem to be effective only in controlling the number of active dislocations and therefore the height of the peak, as it is shown by the Q_m^{-1}-concentration curve of Fig. 47.

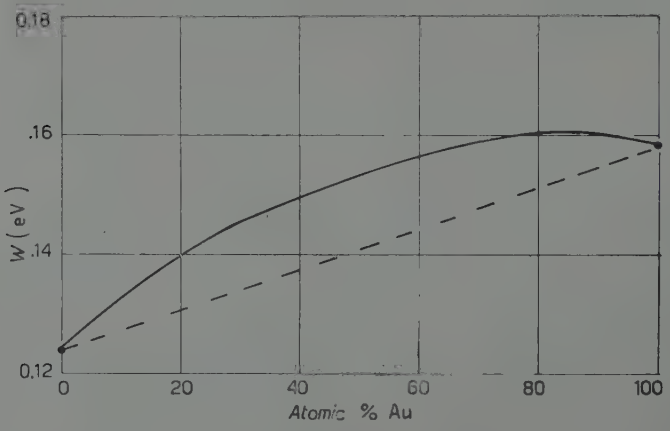


Fig. 49. – Dependence of W on the gold atomic %, in silver-gold alloys. Dotted line, linear dependence: heavy line, assumed dependence [A-18].

The activation energies for the different silver-gold alloys have not yet been measured directly. An estimate for the values of W has been obtained assuming that the energy-concentration curve turns the concavity towards the concentration axis as it happens with the velocity c_e of Fig. 46, and as it is suggested by the negative deviations from the Vegard's law shown by the lattice constant ([16]).

The values of W obtained in this way are given in Fig. 49 as a function of the atomic concentration of gold. It will be interesting to check these values with the direct measurements of the activation energies for the same alloys.

10. – Conclusions.

The theoretical and experimental investigation on the relaxation effect, althought limited to f. c. c. metals, has given substantial contributions to the knowledge of the dislocation behaviour.

The intrinsic character of the mechanism which gives rise to the motion of dislocations near their equilibrium positions has been established beyond doubt by the experiments on the influence of anneal, cold work impurity content and neutron irradiation. This result has suggested a very satisfactory model for dislocation motion: the formation of kink pairs, whose potential barriers and characteristic frequency depend *only* upon the physical parameters of the *single* dislocations, and *not* on their density or on the presence of foreign atoms.

Conversely, the last two parameters have been found very effective in controlling the *number* of active dislocations, as it is shown by the influence of thermal and mechanical treatments, of impurity content and of neutron irradiation on the height of the relaxation peak.

Another general property which has been established is that the volume of the specimens is independent of dislocation motion. It may be added that the dislocations of a given metal have been found to differ between them rather in their characteristic frequencies than in their potential barriers.

The dislocation relaxation in f. c. c. metals has not yet been completely explored. Further information on the effect in single crystals will be helpful in defining the *types* of dislocations responsible for the two dissipation peaks. It would also be interesting to relate the relaxation effect to some direct evidence of the presence of the dislocations, such as can be obtained from X-rays or electron diffraction experiments. Additional infor-

([16]) C. S. Barrett: *Structure of Metals* (London, 1953), p. 222.

mation is required to explain the influence of high temperature anneal upon dissipation and frequency.

However the present review points out the opportunity of focussing the attention on the entirely unexplored field of non-f.c.c. metal crystals. It is felt that a systematic investigation of the effect in this field can be as effective as was the previous investigation on the f. c. c. metals in improving our knowledge of the dynamical behaviour of dislocations.

* * *

The Author is very grateful to Prof. M. NUOVO and Dr. L. VERDINI of the Istituto « O. M. Corbino » for their valid contribution to the present paper.

REFERENCES

A. Experimental.

[1] P. G. BORDONI: *Journ. Acoust. Soc. Am.*, **26**, 495 (1954).
[2] H. E. BOEMMEL: *Phys. Rev.*, **96**, 220 (1954).
[3] H. E. BOEMMEL, W. P. MASON, A. WARNER: *Phys. Rev.*, **99**, 1894 (1955); **102**, 64 (1956).
[4] D. H. NIBLETT and J. WILKS: *Phil. Mag.*, **1**, 415 (1956).
[5] T. S. HUTCHISON and G. J. HUTTON: *Can. Journ. Phys.*, **34**, 1498 (1956).
[6] T. S. HUTCHISON and A. J. FILMER: *Can. Journ. Phys.*, **34**, 159 (1956).
[7] W. P. MASON and H. E. BOEMMEL: *Journ. Acoust. Soc. Am.*, **28**, 930 (1956).
[8] D. H. NIBLETT and J. WILKS: *Phyl. Mag.*, **2**, 1427 (1957).
[9] N. G. EINSPRUCH and R. TRUELL: *Phys. Rev.*, **109**, 652 (1958).
[10] P. G. BORDONI and M. NUOVO: *Effect of Crystal Dislocations upon Vibrations* (Contribution to the Palais de la Science of the Exposition Universelle de Bruxelles, 1958).
[11] H. L. CASWELL: *Journ. Appl. Phys.*, **29**, 1210 (1958).
[12] A. J. FILMER, G. J. HUTTON and T. S. HUTCHISON: *Journ. Appl. Phys.*, **29**, 146 (1958).
[13] V. K. PARÈ: *Experimental and Theoretical Study of low-temperature internal Friction in Copper* (Cornell University, Dept. Eng. Phys., Techn. Rep., n. 4, July 1958).
[14] D. THOMPSON and F. H. GLASS: *Rev. Sci. Instr.*, **29**, 1034 (1958).
[15] P. G. BORDONI, M. NUOVO and L. VERDINI: *Phys. Rev. Lett.*, **2**, 200 (1959).
[16] P. G. BORDONI, M. NUOVO and L. VERDINI: *Nuovo Cimento*, **14**, 273 (1959).
[17] D. O. THOMPSON and D. K. HOLMES: *Journ. Appl. Phys.*, **30**, 525 (1959).
[18] P. G. BORDONI, M. NUOVO and L. VERDINI: *Relaxation of Dislocations in face-centered cubic Metals* (to be published in the Proceedings of the III I.C.A. Congress Stuttgart, September 1959).

788

[19] I. Barducci, M. Nuovo and L. Verdini: *Bordoni Peak in Silver-gold Alloys* (to be published in the Proceedings of the III I.C.A. Congress Stuttgart, September 1959); *Nuovo Cimento*, (to be published).
[20] P. G. Bordoni, M. Nuovo and L. Verdini: *Dislocation Relaxation in Silver, Gold, Palladium and Platinum* (to be published).

B. Theoretical.

[1] P. G. Bordoni: *Ric. Scient.*, **19**, 851 (1949).
[2] W. P. Mason: *Phys. Rev.*, **98**, 1136 (1955); *Bell System Tech. Journ.*, **34**, 903 (1955).
[3] J. Weertman: *Journ. Appl. Phys.*, **26**, 202 (1955).
[4] W. P. Mason: *Journ. Acoust. Soc. Am.*, **27**, 643 (1955).
[5] J. Weertman: *Phys. Rev.*, **101**, 1429 (1956).
[6] W. P. Mason: *Phys. Rev.*, **101**, 1430 (1956).
[7] P. G. Bordoni: Proc. II I.C.A. Congress, 101 (1956).
[8] A. Seeger: *Phil. Mag.*, **1**, 651 (1956).
[9] H. Donth: *Zeits. f. Phys.*, **149**, 111 (1957).
[10] A. Seeger, H. Donth and P. Pfaff: *Discussions Far. Soc.*, **23**, 19 (1957).
[11] W. P. Mason: *Physical Acoustic and the Properties of Solids* (New York, 1958) p. 266-271.

SUPPLEMENTO AL VOLUME XVII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1960
3° Trimestre

Indefinite Metric in Quantum Field Theory.

K. L. NAGY

Istituto di Fisica e Scuola di Perfezionamento in Fisica Nucleare dell'Università - Roma (*)

(ricevuto il 25 Luglio 1960)

CONTENTS. — **1. Introduction.** – **2. State vector space with indefinite metric.** 1. Vector space. 2. Operators. 3. Physical postulates. 4. The η formalism. – **3. Field theoretical examples.** 1. Oscillator. 2. Field with non-relativistic commutator. 3. Free electromagnetic field. 4. Field theories with higher order Lagrangians. 5. Auxiliary fields in quantum field theory. 6. The Lee model. 7. Other fields theories. 8. Heisenberg's theory. 9. Markov's proposal. 10. Froissart's model. – **4. Ghost contributions to propagators.** 1. Contributions from ghost of known type. 2. Contributions found by certain approximation method. – **5. Difficulties connected with the probabilistic interpretation.** 1. Decomposition of the state vector space. 2. Probabilistically interpretable and non-interpretable theories. 3. Proposals for eliminating the difficulties. – **6. Conclusion.**

1. – Introduction.

If in a quantum field theory the commutation relations are different from the usual form, if, for example, the basic commutator

$$[a_k, a_k^*]_{\pm}$$

is not equal to the usual δ but to a non positive definite quantity, then the theory has to be quantized by introducing an indefinite metric.

DIRAC [1] and PAULI [2] were the first to suggest that the introduction of an indefinite metric into a quantum theory might be of considerable inter-

(*) Permanent address: Istitute for Theoretical Physics of the R. Eötvös University, Budapest.

est. Recent field theories with an indefinite metric are constructed for two different reasons. Either it turned out that certain theories can be quantized in a consistent way only by introducing such an indefinite metric, or an indefinite metric was brought in deliberately at the beginning.

In the first group one finds quantum electrodynamics [3, 4], the vector meson field [5], the gravitational field [6], the regularizing auxiliary fields in quantum theory [7] (Pauli-Villars regularization), the field theories with higher order Langrangians [8-12] and the LEE model [13-16]. To the second group belongs HEISENBERG's theory of the elementary particles [17], a proposal of BOGOLJUBOV [18-20], a proposal of MARKOV [21, 22] and a model of FROISSART [23]. The reason why an indefinite metric was introduced in these latter theories (with the exception of the last) is essentially to obtain a divergence free theory; this can be achieved (without extra non-local form-factors) because of the properties of the spectral function ϱ *e.g.* in

$$\Delta'_F = \int \varrho(\varkappa^2)\, \Delta_F(\varkappa^2)\, \mathrm{d}\varkappa^2 ,$$

only by introducing an indefinite metric. On the other hand, in a theory having a higher order Lagrangian, or in the Pauli-Villars regularization method, the infinities may disappear just because of this principle.

Quantum field theories with an indefinite metric, however, may, and in most cases actually do, lead to difficulties in connection with the probabilistic interpretation.

This article intends to give a short review of the problems touched above.

2. – State vector space with indefinite metric.

This part represents a simple mathematical introduction to the properties of a vector space with an indefinite metric. There are several methods of treating the problem [15, 26]. Our treatment is similar to that given by GUPTA [24] and PANDIT [25], which seems the most simple and convenient for our purpose. As the vector spaces of the field theories have an infinite number of dimensions, we meet here extra difficulties [25]. Since, however, an indefinite « physical » scalar product may be introduced also over a Hilbert space (cfr. Sect. 2·4), it seeems that these difficulties may be treated along the usual lines. Therefore these problems are not touched here.

2·1. *Vector space.* – The space $\mathscr{H}$ consisting of the elements $|\hat{A}\rangle$, $|\hat{B}\rangle$, ... is a Hilbert space if

1) it is linear;

2) there corresponds to each $|\hat{A}\rangle$ and $|\hat{B}\rangle$ a (complex) number with the properties:

(i) $\langle\hat{A}|\hat{B}\rangle = \{\langle\hat{B}|\hat{A}\rangle\}^*$,

(ii) $\langle\hat{A}|\{\alpha|\hat{B}\rangle + \beta|\hat{C}\rangle\} = \alpha\langle\hat{A}|\hat{B}\rangle + \beta\langle\hat{A}|\hat{C}\rangle$,

(iii) $\langle\hat{A}|\hat{A}\rangle \geqslant 0$,

(iv) $\langle\hat{A}|\hat{A}\rangle = 0$ if, and only if, $|\hat{A}\rangle = 0$.

The other properties required for the definition of a Hilbert space will not concern us here.

The space $\mathscr{H}$ consisting of the elements $|A\rangle$, $|B\rangle$, ... is a vector space with indefinite metric, if (iii) and (iv) are violated, *i.e.* if

a) it is linear;

b) there corresponds to each $|A\rangle$ and $|B\rangle$ a (complex) number $\langle A|B\rangle$ with the properties (i) and (ii).

$\langle A|B\rangle$ is called the scalar product of $|A\rangle$ and $|B\rangle$. If $\langle A|B\rangle = 0$, $|A\rangle$ and $|B\rangle$ are orthogonal. $\langle A|A\rangle$ is the norm of $|A\rangle$. From (i) $\langle A|A\rangle$ is a real number. If $\langle A|A\rangle \neq 0$, $|A\rangle$ can be always normalized to ± 1.

If (iii) and (iv) are violated in such a way that $\langle A|A\rangle \geqslant 0$, (or $\langle A|A\rangle \leqslant 0$) but $\langle A|A\rangle = 0$ also for an $|A\rangle \neq 0$, $\mathscr{H}$ is a vector space with semi-definite metric; if there exists at least one $|A\rangle$ with $\langle A|A\rangle = 0$, orthogonal to all the vectors, $\mathscr{H}$ is called a vector space with degenerate metric. If the metric is not degenerate, there exists an orthonormal basis $|i\rangle$, ... with

$$\langle i|j\rangle = N_i\delta_{ij}, \qquad N_i = \pm 1$$

spanning the space $\mathscr{H}$. We do not deal in the following with the degenerate case.

2·2. *Operators.* – An operator P, as usual, has the property of transforming the vectors of the space among themselves. P is linear, if

$$P\{\alpha|A\rangle + \beta|B\rangle\} = \alpha P|A\rangle + \beta P|B\rangle.$$

P^* is the hermitian conjugate of P, if for each $|A\rangle$, $|B\rangle \in \mathscr{H}$

$$\langle A|P^*|B\rangle = \{\langle B|P|A\rangle\}^*.$$

Thus it follows that $(P^*)^* = P$ and

$$\langle A|PQ|B\rangle = \{\langle B|Q^*P^*|A\rangle\}^*, \quad \text{etc.}$$

If $P|p\rangle = p|p\rangle$, where p is a number, $|p\rangle$ is an eigenvector and p is an eigenvalue of P. The mean value $\bar{P}$ in the state $|A\rangle$ of P is defined by $\bar{P} = \langle A|P|A\rangle$.

If $P = P^*$, P is said to be hermitian. In this case it simply follows that:

(2.1) if $P|p\rangle = p|p\rangle$, $\qquad (p - p^*)\langle p|p\rangle = 0$,

(2.2) if $P|p_1\rangle = p_1|p_1\rangle$ and $P|p_2\rangle = p_2|p_2\rangle$, $\quad (p_1^* - p_2)\langle p_1|p_2\rangle = 0$,

(2.3) $\bar{P}$ is a real number.

Since here $\langle p|p\rangle$ may be also zero, p is not necessarily real.

An operator satisfying $UU^* = U^*U = 1$ is called a unitary operator. The unitary transformation

$$|A\rangle' = U|A\rangle, \qquad P' = UPU^{-1}$$

has the following properties:

if $P|p\rangle \quad p|p\rangle$ then $P'|p\rangle' = p|p\rangle'$ and $'\langle p|P'|p\rangle' = \langle p|P|p\rangle$.

If $|i\rangle, \ldots$ is an orthogonal basis, and we define the diadic product $|A\rangle\langle B|$ of state vectors in the usual way, the unit operator is

(2.4) $$1 = \sum |i\rangle N_i \langle i|.$$

This allows us to represent $|A\rangle$ and P by the matrices

(2.5) $$\begin{cases} c_i = N_i\langle i|A\rangle, \\ P_{ij} = N_i\langle i|P|j\rangle. \end{cases}$$

P_{ij} fulfils the matrix multiplication rule. Indeed

$$(PQ)_{nm} = N_n\langle n|PQ|m\rangle = N_n \sum \langle n|P|i\rangle N_i\langle i|Q|m\rangle = \sum P_{ni}Q_{im}.$$

From (2.5)

$$P_{ij}^* = N_i\langle j|P^*|i\rangle.$$

Thus if $P = P^*$

$$P_{ij}^* = N_i N_j P_{ji} ,$$

but if $UU^* = 1$, $\sum U_{ki}^* U_{il} = \delta_{kl}$ as usual. With the representation (2.5), the eigenvalue equation

$$P|p\rangle = p|p\rangle$$

has the well known matrix form, since

$$\sum P|i\rangle N_i\langle i|p\rangle = p \sum |i\rangle N_i\langle i|p\rangle , \qquad \sum \langle j|P|i\rangle c_i = pN_j c_j ,$$

thus

$$\sum P_{ij} c_i = pc_j .$$

The most important question is: when do the eigenvectors of an operator P form a complete system? In a finite dimensional (n) vector space this can be answered as follows:

Let us define the minumum polynomial $m(P)$ of P as the polynomial of least possible degree for which

$$m(P) \equiv 0 .$$

Then the necessary and sufficient condition that the eigenvectors of an operator P form a complete system is that its minimum polynomial should be written as

$$m(P) = (P - p_1) \ldots (P - p_k) , \tag{2.6}$$

where $p_i \neq p_j$ if $i \neq j$. The p_i are then eigenvalues.

At least one polynomial for which $M(P) \equiv 0$ always exists, is the characteristic polynomial of P. It can be written

$$M(P) = (P - p_1)^{\alpha_1} \ldots (P - p_k)^{\alpha_k} , \qquad \alpha_1 + \ldots + \alpha_k = n .$$

Thus the minimum polynomial always has the form

$$m(P) = (P - p_1)^{\beta_1} \ldots (P - p_k)^{\beta_k} , \qquad \beta_i \leqslant \alpha_i .$$

The question arises: do the eigenvectors of a hermitian operator form a complete system? If this were not the case then at least one $\beta_i = k > 1$. Then there exists at least one state vector $|A\rangle \neq 0$ such that

$$(P - p_i)^k |A\rangle = 0 , \quad \text{but} \quad (P - p_i)^{k-1} |A\rangle = |B\rangle \neq 0 . \tag{2.7}$$

$|B\rangle$ satisfies the eigenvalue equation

$$(P - p_i)|B\rangle = 0 .$$

There are two cases: p_i is either complex or real. In the first case because of (2.1)

$$\langle B|B\rangle = 0 .$$

In the second case

$$\langle B|B\rangle = \langle A|(P - p_i)^{2(k-1)}|A\rangle = 0 , \quad \text{since } 2(k-1) \geqslant k .$$

In a Hilbert space, if the norm of $|B\rangle$ vanishes, then also $|B\rangle = 0$, thus in a Hilbert space the eigenfunctions of a hermitian operator always form a complete system. This is no more true in the case of an indefinite metric. The necessary condition that the eigenfunctions of a hermitian operator do not form a complete system is that at least one eigenvector with zero norm should exist.

The above theorem provides a method of completing the vector space. Suppose we have found all the eigenvectors of a hermitian operator. We consider only those which have zero norm. Be $|B\rangle$ one of them with the eigenvalue p. Then one has to decide whether there exists or not a vector $|A\rangle \neq 0$ satisfying

$$(P - p)|A\rangle = |B\rangle . \tag{2.8}$$

If such vector exists, it has to be added for completing the system. Repeating the procedure, if necessary, the vector space can be completed.

In a vector space with non degenerate metric, the following theorem holds: if there is at least one eigenvector among all the other eigenvectors of a hermitian operator which has zero norm, orthogonal to the other eigenvectors, then at least one $|A\rangle$ satisfying (2.8) necessarily exists.

2·3. *Physical postulates.* – The above vector space with indefinite metric $\mathscr{H}$ has now to be connected with physical systems. This can be done by the following postulates:

(i) the state of a system is described by an $|A\rangle \in \mathscr{H}$ at each instant of time.

(ii) the physical quantities are represented by linear operators operating in $\mathscr{H}$;

(iii) if in a state $|A\rangle$ P is measured, the average result of the measurements is $\bar{P} = \langle A|P|A\rangle$.

These postulates imply the following consequences:

(*a*) since $\overline{P}$ must be real, the physical quantities have to be represented by hermitian operators;

(*b*) if $\langle A|A\rangle = 0$, but $\langle A|P|A\rangle \neq 0$, $\overline{P}$ is undetermined, since $|A\rangle$ can be multiplied by an arbitrary number;

(*c*) let $f(P)$ be any arbitrary function of P, which can be developed in powers of P, and form $P|p_i\rangle = p_i|p_i\rangle$ a complete system of orthonormal vectors. Then

$$\langle A|f(P)|A\rangle = \sum_{m,n} c_m^* c_n f(p_n)\langle m|n\rangle = \sum_m N_m c_m^* c_m f(p_m) .$$

But the average result of the measurement of $f(P)$ is

$$\overline{P} = \sum W_m f(p_m)$$

where W_m is the probability of measuring p_m. Since f is arbitrary

$$W_m = N_m c_m^* c_m , \qquad N_m = \pm 1 .$$

From (*b*) and (*c*) it follows that there are difficulties with states of zero, or negative norm. These problems require further special considerations.

2·4. *The η-formalism.* – The formalism of the previous section is often replaced in the literature by a method in which a metric operator η is defined in a Hilbert space. We here draw the connection between these two treatments.

Let us construct a Hilbert space $\hat{\mathscr{H}}$ so that there be a one to one correspondence between the state vectors $|A\rangle$ ... of $\mathscr{H}$ and $|\hat{A}\rangle$ ... of the Hilbert space $\hat{\mathscr{H}}$, with the property: if $|A\rangle \to |\hat{A}\rangle$ and $|B\rangle \to |\hat{B}\rangle$ then

$$\alpha|A\rangle + \beta|B\rangle \to \alpha|\hat{A}\rangle + \beta|\hat{B}\rangle ; \tag{2.9}$$

for each $|A\rangle$ and $|B\rangle$. Because of this correspondance, the bilinear form $\langle A|B\rangle$ defines an operator η in the Hilbert space:

$$\langle A|B\rangle = \langle \hat{A}|\eta|\hat{B}\rangle . \tag{2.10}$$

The operator $\hat{P}$ corresponding to the linear operator P is defined by

$$\langle A|P|B\rangle = \langle \hat{A}|\eta\hat{P}|\hat{B}\rangle . \tag{2.11}$$

The hermitian conjugate $\hat{P}$ of $\hat{P}$ in $\hat{\mathscr{H}}$ fulfils as usual

$$\{\langle\hat{A}|\hat{P}|\hat{B}\rangle\}^* = \langle\hat{B}|\hat{P}\ |\hat{A}\rangle\,.$$

Since $\langle A|A\rangle = \langle\hat{A}|\eta|\hat{A}\rangle$ is real, $\eta^+ = \eta$. η is called the metric operator. η always operates in $\hat{\mathscr{H}}$, therefore here the roof ($\hat{}$) is omitted. From (2.11) it follows that

$$\overline{P} = \langle A|P|A\rangle = \langle\hat{A}|\eta\hat{P}|\hat{A}\rangle\,. \tag{2.12}$$

Defining the operators $\hat{P}$ and $\hat{P}^+$, we can easily express the operator $\hat{P}^*$ in terms of $\hat{P}^+$. Indeed

$$\langle A|P^*|B\rangle = \{\langle B|P|A\rangle\}^* = \{\langle\hat{B}|\eta\hat{P}|\hat{A}\rangle\}^* = \langle\hat{A}|\hat{P}^+\eta|\hat{B}\rangle = \langle\hat{A}|\eta\eta^{-1}\hat{P}^+\eta|\hat{B}\rangle.$$

Since $|A\rangle$ and $|B\rangle$ are arbitrary, from (2.11) one finds that the operator corresponding to P^* is

$$\hat{P}^* = \eta^{-1}\,\hat{P}^+\eta\,. \tag{2.13}$$

Therefore, if P is hermitian in $\mathscr{H}$ ($P = P^*$), it follows from (2.13) that $\hat{P} = \eta^{-1}\,\hat{P}^+\eta$ thus $\hat{P} = \hat{P}^+$, *i.e.* $\hat{P}$ is hermitian in $\hat{\mathscr{H}}$ only if $[\hat{P}, \eta] = 0$, and reversely.

For spanning $\mathscr{H}$ we can use an orthogonal system $|i\rangle, \ldots, \langle i|j\rangle = N_i\delta_{ij}$. We choose one arbitrarily, but fix it once and for all afterwards, and construct the Hilber space $\hat{\mathscr{H}}$ so that it be spanned by the orthonormal system $|\hat{i}\rangle, \ldots, \langle\hat{i}|\hat{j}\rangle = \delta_{ij}$.

From (2.10) then

$$N_i\delta_{ij} = \langle\hat{i}|\eta|\hat{j}\rangle\,, \qquad \eta|j\rangle = N_j|j\rangle\,, \tag{2.14}$$

i.e. in this representation η is diagonal having the eigenvalues $N_j = \pm 1$.

Introducing in $\mathscr{H}$ a new orthonormal system

$$|i\rangle' = U|i\rangle\,, \qquad UU^* = 1\,, \qquad '\langle i|j\rangle' = N_i\delta_{ij}\,,$$

the system

$$|\hat{i}\rangle' = \hat{U}|\hat{i}\rangle$$

in general is not orthogonal, since according to (2.11) and (2.13) $\hat{U}^*\hat{U} = \eta^{-1}\hat{U}^+\eta\hat{U}$ and not $\hat{U}^+\hat{U} = 1$. Having started from the system $|i\rangle'$ and $|\hat{i}\rangle'$, $|\hat{i}\rangle$ would not be orthogonal.

In Sect. 3·1, we shall show a simple example for this formalism, but we do not use it in the following.

3. – Field theoretical examples.

In this part, field theoretical examples, — mainly those which were quoted in the introduction — are treated. We discuss here only the method of obtaining a theory with indefinite metric, while the difficulties like those connected with the probabilistic interpretation will be treated later.

3·1. *Oscillator.* – Since free fields in quantum theory are equivalent to an assembly of harmonic oscillators, we consider first a simple harmonic oscillator leading to definite or indefinite metric. Thus later we will not need to discuss similar problems again and again.

3·1.1. Normal oscillator. In the case of a simple harmonic oscillator, the operators satisfying the commutation relation

$$aa^* - a^*a = 1 \tag{3.1}$$

forms a complete set of operators in the sense of Schur's lemma: an operator commuting with them must be a multiple of the identity operator. Each physical quantity can be expressed be means of them. For example, the Hamiltonian is

$$H = Ea^*a\,. \tag{3.2}$$

The state vector space is built up by introducing the vacuum state with minimal energy, having the properties

$$a|0\rangle = 0\,, \qquad \langle 0|0\rangle = 1\,, \tag{3.3}$$

and then the orthogonal system is

$$|n\rangle = (n!)^{-\frac{1}{2}}(a^*)^n|0\rangle\,.$$

The scalar product of $|m\rangle$ and $|n\rangle$, $\langle m|n\rangle = (m!n!)^{-\frac{1}{2}}\langle 0|a^m a^{*n}|0\rangle = \delta_{mn}$ where in the last equality (3.1) and (3.3) were used. Of course a^* is the hermitian conjugate of a, and

$$H|n\rangle = nE|n\rangle\,.$$

The state vector space is a Hilbert space, and hermitian conjugation etc. are understood conventionally.

3'1.2. «Abnormal» oscillator. Instead of (3.1) and (3.2) in this case *e.g.* a Lagrangian formalism gives

$$aa^* - a^*a = -1\,, \tag{3.4}$$

$$H = -Ea^*a\,. \tag{3.5}$$

Here, obviously, we have two possibilities. The vacuum state may be defined by

$$a^*|0\rangle = 0\,, \qquad \langle 0|0\rangle = 1\,,$$

and the orthogonal system as

$$|n\rangle = (n!)^{-\frac{1}{2}}(a)^n|0\rangle\,.$$

In this case we obtain a Hilbert space, but the energy becomes negative. This possibility for the field theories treated later is excluded also because of other reasons (for quantum electrodynamics see for example [27], for the Lee model the conservation of the fermion number forbids this teatment etc.), therefore one has to choose the second possibility resulting in an indefinite metric.

Then the vacuum state and the orthonormal system is defined by

$$\begin{cases} a|0\rangle = 0\,, \qquad \langle 0|0\rangle = 1\,. \\ |n\rangle = (n!)^{-\frac{1}{2}}(a^*)^n|0\rangle\,, \end{cases} \tag{3.6}$$

thus from (3.4)

$$\langle m|n\rangle = (m!n!)^{-\frac{1}{2}}\langle 0|a^m a^{*n}|0\rangle = (-1)^n\delta_{mn}\,, \tag{3.7}$$

i.e. a state vector space with indefinite metric. a^* is the hermitian conjugate of a. Here also

$$H|n\rangle = nE|n\rangle\,, \tag{3.7'}$$

but

$$\bar{H}_n = \langle n|H|n\rangle = (-1)^n nE\,.$$

H is obviously hermitian. The unit operator is here

$$1 = \sum |n\rangle(-1)^n\langle n|\,,$$

therefore the state vector $|A\rangle$ is represented by

$$|A\rangle = \sum c_n |n\rangle , \qquad c_n = (-1)^n \langle n|A\rangle .$$

According to (3.4) and (3.7)

$$a|n\rangle = -\sqrt{n}\,|n-1\rangle , \qquad a^*|n\rangle = \sqrt{n+1}\,|n+1\rangle , \tag{3.8}$$

and thus from (2.5)

$$a_{mn} = -\sqrt{n}\,\delta_{m+1,n} , \qquad a^*_{mn} = \sqrt{n+1}\,\delta_{m,m+1} .$$

H is diagonal in this representation, and therefore its « minimum polynomial » is

$$m(H) = \prod_{n=0}^{\infty} (H - nE) .$$

We now show how this oscillator can be treated by means of the η formalism. Denote by $|\hat{n}\rangle$ ... the complete orthonormal system of state vectors in a Hilbert space corresponding to $|n\rangle$, ...

$$\langle \hat{m}|\hat{n}\rangle = \delta_{mn} . \tag{3.9}$$

Then according to (2.14)

$$\langle m|n\rangle = (-1)^n \delta_{mn} = \langle \hat{m}|\eta|\hat{n}\rangle , \qquad \eta|\hat{n}\rangle = (-1)^n |n\rangle . \tag{3.10}$$

Since from (3.7) and (3.8)

$$\langle m|a^*|n\rangle = (-1)^m \sqrt{n+1}\,\delta_{m,n+1} ,$$

it then follows from (2.12)

$$\hat{a}^*|\hat{n}\rangle = \sqrt{n+1}\,|\widehat{n+1}\rangle , \tag{3.11}$$

η being diagonal. Applying η on both sides we find

$$\eta\hat{a}^*\eta^{-1}\eta|\hat{n}\rangle = \eta\hat{a}^*\eta^{-1}(-1)^n|\hat{n}\rangle = \sqrt{n+1}(-1)^{n+1}|\widehat{n+1}\rangle = (-1)^{n+1}\hat{a}^*|\hat{n}\rangle ,$$

from which

$$\eta\hat{a}^*\eta^{-1} = -\hat{a}^* . \tag{3.12}$$

From (2.13)

$$\hat{a} = \eta^{-1}\hat{a}^{*+}\eta . \tag{3.13}$$

Denoting $\hat{a}^* = A^+$ and $\hat{a}^{*+} = A$, according to (3.4), (3.12) and (3.13) (cfr. 2.11)

800

one obtains

$$(3.14)\qquad \begin{cases} -1 = \hat{a}\,\hat{a}^* - \hat{a}^*\hat{a} = \eta^{-1}A\eta A^+ - A^+\eta^{-1}A\eta = -\eta^{-1}(AA^+ - A^+A)\eta\,, \\ AA^+ - A^+A = 1\,. \end{cases}$$

From the operators A we can construct now the system $|\hat{n}\rangle$ ... defining

$$(3.15)\qquad A|\hat{0}\rangle = 0\,, \qquad \langle\hat{0}|\hat{0}\rangle = 1\,, \qquad |n\rangle = (n!)^{-\frac{1}{2}}(A^+)^n|\hat{0}\rangle\,,$$

since from (3.14) $A^+A|\hat{n}\rangle = n|\hat{n}\rangle$,

$$(3.16)\qquad \eta = (-1)^{A^+A}\,.$$

It is easy to see that with this choice all necessary formulae are satisfied and

$$(3.17)\qquad H = EA^+A\,.$$

Having started from (3.14)-(3.17), but defined the « physical » scalar product and mean value according to (2.10) and (2.12), we would get the same physical results as without η. This is the method of the η-formalism.

3·2. *Field with non-relativistc commutator.* – Here a free Dirac field with an equal-time commutator not of δ-type will be treated as an extremely simple example of a field theory with an indefinite metric.

Let the field equation and commutation relations be

$$(\gamma_\mu\partial_\mu + k)\psi = 0$$

$$\{\psi_\alpha(\boldsymbol{x}, t), \psi^*_\beta(\boldsymbol{x}', t)\} = \delta_{\alpha\beta}f(\boldsymbol{x} - \boldsymbol{x}')\,;$$

other anticommutator as usual.

Here it is supposed

$$f(\boldsymbol{x}) = f^*(\boldsymbol{x})\,, \qquad f(\boldsymbol{x}) = f(-\boldsymbol{x})\,.$$

Then the Fourier transforms $f(\boldsymbol{p})$ of f fulfils:

$$f(\boldsymbol{p}) = \int f(\boldsymbol{x})\exp[i\boldsymbol{p}\boldsymbol{x}]\,\mathrm{d}\boldsymbol{x}\,, \qquad f(\boldsymbol{p}) = f^*(\boldsymbol{p})\,, \qquad f(\boldsymbol{p}) = f(-\boldsymbol{p})\,.$$

ψ satisfies the field equation and the commutation relation if

$$\psi(x) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int\sqrt{|f(\boldsymbol{p})|}\{a(\boldsymbol{p}, s)\exp[ipx]u(\boldsymbol{p}, s) + b^*(\boldsymbol{p}, s)\exp[-ipx]v(\boldsymbol{p}, s)\}\mathrm{d}\boldsymbol{p}\,,$$

$$\overline{\psi}(x) = \frac{1}{(2\pi)^{\frac{3}{2}}}\int\sqrt{|f(\boldsymbol{p})|}\{b(\boldsymbol{p}, s)\exp[ipx]\,\overline{v}(\boldsymbol{p}, s) + a^*(\boldsymbol{p}, s)\exp[-ipx]\,\overline{u}(\boldsymbol{p}, s)\}\mathrm{d}\boldsymbol{p}\,,$$

and

$$\{a(\boldsymbol{p}, s), a^*(\boldsymbol{p}', s')\} = \{b(\boldsymbol{p}, s), b^*(\boldsymbol{p}', s')\} = \operatorname{sign} f(\boldsymbol{p})\,\delta(\boldsymbol{p} - \boldsymbol{p}')\,\delta_{ss'}\,.$$

P_μ which is defined by $[P_\mu, \psi] = i\partial_\mu \psi$ is

$$P_\mu = \frac{1}{(2\pi)^3}\int p_\mu \operatorname{sign} f(\boldsymbol{p})\,\{a^*(\boldsymbol{p}, s)a(\boldsymbol{p}, s) + b^*(\boldsymbol{p}, s)b(\boldsymbol{p}, s)\}\,\mathrm{d}\boldsymbol{p}\,.$$

(If $f(\boldsymbol{x}) \neq \delta(\boldsymbol{x})$ the Lagrangian formalism does not work.) For momenta with sign $f(\boldsymbol{p}) = -1$, we have «abnormal» oscillators and one thus obtains a field theory with indefinite metric. Obiously ψ^* is the hermitian conjugate of ψ and $P_i^* = P_i$, $P_0^* = P_0$.

3·3. *Free electromagnetic field.* – The Lagrangian

$$L = -\tfrac{1}{2}\,\partial_\mu A_\nu\,\partial_\mu A_\nu$$

give in the usual way

$$\Box A_\mu = 0$$

$$[A_\mu(x), A_\nu(x')] = i\delta_{\mu\nu} D(x - x')\,,$$

$$P_\nu = \int_\sigma (\partial_\mu A_\varrho\,\partial_\nu A_\varrho - \tfrac{1}{2}\,\delta_{\mu\nu}\partial_\varrho A_\sigma\,\partial_\varrho A_\sigma)\,\mathrm{d}\sigma_\mu\,.$$

These field equations and commutation relations are satisfied if

$$A_\mu(x) = \sum \sqrt{\frac{1}{2\,|k|}}\,(a(k, \lambda)\exp[ikx] + \tilde{a}(k, \lambda)\exp[-ikx])\,e_\mu(k, \lambda)\,,$$

with

$$e_\mu(k, \lambda)\,e_\mu(k, \lambda') = \delta_{\lambda\lambda'}\,, \qquad \sum_\lambda e_\mu(k, \lambda)\,e_\nu(k, \lambda) = \delta_{\mu\nu}\,,$$

$$e_\mu(k, \lambda)k_\mu = \begin{cases} = 0\,, & \lambda = 1, 2 \\ = |k|\,, & \lambda = 3 \\ = i\,|k|\,, & \lambda = 4\,. \end{cases}$$

$$\tilde{a}(k, i) = a^*(k, i) \quad \text{for } i = 1, 2, 3; \quad a(k, 4) = -ia(k, 0); \quad \tilde{a}(k, 4) = -ia^*(k, 0)\,,$$

$$[a(k, i),\ a^*(k', i')] = \delta_{kk'}\delta_{ii'}\,,$$

$$[a(k, 0),\ a^*(k', 0)] = -\delta_{kk'}\,.$$

P_μ expressed by the a is

$$P_\mu = \sum k_\mu \Big\{\sum_{i=1}^{3} a^*(k, i)\,a(k, i) - a^*(k, 0)a(k, 0)\Big\}\,.$$

802

Because of the commutator of $a(k, 0)$ one obtains a field theory with indefinite metric. The state vectors are of the type (3.6): $|N_1, N_2, N_3, N_0\rangle$ with the norm

$$\langle N_1, N_2, N_3, N_0 | N_1', N_2', N_3', N_0' \rangle = (-1)^{N_0} \delta_{NN'} .$$

Obviously $A_i = A_i^*$, $A_0 = A_0^*$.

In order that Maxwell's equations are satisfied as mean value equations, one has to prescribe the Lorentz condition by postulating that only those state vectors are allowed to describe physical states for which

$$\partial_\mu A_\mu^{(+)} |L\rangle = 0 .$$

For a fixed $\boldsymbol{k} \| z$ this condition allows for N_3, N_0 the following set of states (Lorentz set):

$$|0\rangle = |0, 0\rangle ,$$

$$|1\rangle = |1, 0\rangle - |0, 1\rangle ,$$

$$. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .$$

Thus

$$|L\rangle = \sum c_n |n\rangle .$$

Since

$$\langle m | n \rangle = 0 \quad \text{if} \quad m \neq n \neq 0 , \qquad \langle 0 | 0 \rangle = 1 ,$$

$$\langle L | L \rangle = c_0^* c_0 \geqslant 0 ,$$

thus in the Lorentz set $|L\rangle$ one does never meet negative probabilities, and

$$\langle L | P_\mu | L \rangle = \langle L | \sum_{k, \lambda = 1}^{2} k_\mu a^*(k, \lambda) a(k, \lambda) | L \rangle .$$

The subspace of $\mathscr{H}$ spanned by $|L\rangle$ is a vector space with degenerate metric, the whole space $\mathscr{H}$, however, is non-degenerate.

3·4. *Field theories with higher order Lagrangians.* – For the sake of simplicity we treat here the case of a simple scalar field φ satisfying the equation

$$\left(1 - \frac{\Box}{\mu^2}\right) \Box \varphi = - \varrho ,$$

where ϱ is the source function (operator). The Lagrangian density corre-

sponding to the field equation is

$$L = -\frac{1}{2}\,\partial_\nu\varphi\left(1-\frac{\Box}{\mu^2}\right)\partial_\nu\varphi - \varrho\varphi\,.$$

Introducing

$$\varphi^{(0)} = \left(1-\frac{\Box}{\mu^2}\right)\varphi\,,$$
$$\varphi^{(1)} = -\frac{\Box}{\mu^2}\,\varphi\,, \qquad \varphi = \varphi^{(0)} - \varphi^{(1)},$$

the Lagrangian density L is equal up to a four divergence to the Lagrangian density L'

$$L' = -\tfrac{1}{2}\,\{\partial_\nu\varphi^{(0)}\partial_\nu\varphi^{(0)} - \partial_\nu\varphi^{(1)}\partial_\nu\varphi^{(1)} - \mu^2\varphi^{(1)2}\} - \varrho(\varphi^{(0)} - \varphi^{(1)})\,.$$

From this one gets the field equations and commutation relations

$$\Box\varphi^{(0)} = -\varrho\,, \qquad [\varphi^{(0)}(x),\ \varphi^{(0)}(x')] = iD(x-x')\,,$$
$$(x-x')^2 \geqslant 0\,.$$

$$(\Box - \mu^2)\varphi^{(1)} = -\varrho\,, \qquad [\varphi^{(1)}(x),\ \varphi^{(1)}(x')] = -i\,\Delta(x-x')\,,$$

If we span the state vector space by free particle states, then, because of the negative sign in the commutator of $\varphi^{(1)}$ we obtain a field theory with indefinite metric:

$$\langle N_0, N_1 | N_0', N_1'\rangle = (-1)^{N_1}\delta_{N,N'}\,.$$

$\varphi^{(0)}$, $\varphi^{(1)}$ are hermitian. Field theories with higher order Lagrangian automatically lead to regularized propagators. Indeed for example here

$$[\varphi(x),\ \varphi(x')] = i\big(D(x-x') - \Delta(x-x')\big)\,,$$

from which the δ-like singularity already has dropped out.

3·5. *Auxiliary fields in quantum field theory.* – The Pauli-Villars regularization method works with appropriately regularized propagators. This type of propagator appears in the S-matrix, if one considers the interaction Hamiltonian

$$H(x) = g\bar{\psi}(x)\psi(x)\varphi(x)$$

e.g. for the sake of simplicity, where

$$\psi(x) = \psi_0(x) + \sum c_n \psi_n(x) .$$
$$\varphi(x) = \varphi_0(x) + \sum b_n \varphi_n(x) .$$

b_n is real. In order to obtain a finite theory, two auxiliary fermion fields, and one auxiliary boson field are sufficient [20]. We prescribe the commutation relations

$$\{\psi_m(x), \bar{\psi}_n(x')\} = -i\,\varepsilon_n S(x - x'; K_n^2)\,\delta_{nm} , \qquad \varepsilon_0 = 1 ,$$
$$[\varphi_n(x), \varphi_m(x')] = i\,\varrho_n \Delta(x - x'; \mu_n^2)\,\delta_{nm} , \qquad \varrho_0 = 1 ,$$

from which

$$\{\psi(x), (\bar{\psi}x')\} = -i \sum \varepsilon_n c_n^* c_n S(x - x'; K_n^2) ,$$
$$[\varphi(x), \varphi(x')] = i \sum \varrho_n b_n^2 \Delta(x - x'; \mu_n^2) .$$

For ε, ϱ, c, b the Pauli Villars conditions are prescribed

$$\sum \varepsilon_n c_n^* c_n = 0 , \qquad \sum \varepsilon_n K_n^2 c_n^* c_n = 0 , \ \ldots$$
$$\sum \varrho_n b_n^2 = 0 , \qquad \sum \varrho_n \mu_n^2 b_n^2 = 0 , \ \ldots$$

Because of these conditions, we have necessarily negative values for ε and ϱ, and therefore, considering the commutators, it is a theory with indefinite metric.

We mention here that if there is at least one auxiliary fermion field, the fermion self energy becomes linearly divergent [20], since one has to substitute different masses also into the numerator of a fermion propagator.

3·6. *The Lee model.* – PAULI and KÄLLEN were the first to stress [14] that the Lee model in its original form [13] without cut-off leads to a contradiction. Indeed the theory gives

$$|c|^2 = \frac{g^2}{g_0^2} = 1 - g^2 \int \frac{k^2\,\mathrm{d}k}{2\omega(\omega - E)} = -\infty . \qquad \text{for } g \text{ finite.} \tag{3.18}$$

Here g_0 and g are the bare and observed coupling constants, $|c|^2$ is the probability of finding one bare V particle in a dressed V-particle state. They have shown that by introducing an indefinite metric a consistent field theory can be developed. Here we present the model using the notations of HEISENBERG [15].

The Lee model starts from the Hamiltonian

$$H = -m_{\mathrm{V}}\int \psi_{\mathrm{V}}^{*}\psi_{\mathrm{V}}\,\mathrm{d}\boldsymbol{p} + \int \omega(k)\,a^{*}(\boldsymbol{k})\,a(\boldsymbol{k})\,\mathrm{d}\boldsymbol{k} - \tag{3.19}$$

$$- \frac{g_0}{\sqrt{4\pi}}\int \frac{1}{\sqrt{2\omega}}\,\delta(\boldsymbol{p}-\boldsymbol{q}-\boldsymbol{k})\left\{\psi_{\mathrm{N}}^{*}(\boldsymbol{q})\,\psi_{\mathrm{V}}(\boldsymbol{p})\,a^{*}(\boldsymbol{k}) - \psi_{\mathrm{V}}^{*}(\boldsymbol{p})\,\psi_{\mathrm{N}}(\boldsymbol{q})\,a(\boldsymbol{k})\right\}\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{q}\,\mathrm{d}\boldsymbol{k}\,.$$

Here ψ_{V}, ψ_{N} and a stand for the field operators of the V, N (fermions) and θ (boson) particles, m_{V} is the bare V particle mass, we put $m_{\mathrm{N}} = 0$. g_0 is the bare coupling constant, which is pure imaginary.

(3.19) allows the transition $\mathrm{V} \rightleftarrows \mathrm{N} + \theta$ and $N_1 = N_{\mathrm{V}} + N_{\mathrm{N}}$, $N_2 = N_{\mathrm{V}} + N_{\theta}$ are conserved. The commutation relations are

$$\left\{\psi_{\mathrm{V}}(\boldsymbol{p}),\, \psi_{\mathrm{V}}^{*}(\boldsymbol{p}')\right\} = -\,\delta(\boldsymbol{p}-\boldsymbol{p}')\,, \qquad \left\{\psi_{\mathrm{N}}(\boldsymbol{p}),\, \psi_{\mathrm{N}}^{*}(\boldsymbol{p}')\right\} = \delta(\boldsymbol{p}-\boldsymbol{p}')\,,$$

$$[a(\boldsymbol{k}),\, a^{*}(\boldsymbol{k}')] = \delta(\boldsymbol{k}-\boldsymbol{k}')\,,$$

the others as usual. The vacuum state is defined by

$$\psi_{\mathrm{V}}|0\rangle = \psi_{\mathrm{N}}|0\rangle = a|0\rangle = 0\,, \qquad \langle 0|0\rangle = 1\,, \tag{3.20}$$

and the bare particle states by applying ψ_{V}^{*}, ψ_{N}^{*} and a^{*} on it. Because of the negative sign in the ψ_{V} anticommutator, one has here a field theory with indefinite metric:

$$\langle N_{\mathrm{V}}, N_{\mathrm{N}}, N_{\theta}|N_{\mathrm{V}}', N_{\mathrm{N}}', N_{\theta}'\rangle = (-1)^{N_{\mathrm{V}}}\delta_{\mathrm{NN}'}\,.$$

The operators with stars are the hermitian conjugate of those without and

$$H = H^{*}\,.$$

In the following, we consider the real particle states.
The real vacuum and the real N particle states

$$H|0\rangle = 0|0\rangle\,,$$

$$H|N\rangle = 0|N\rangle$$

are equal to the corresponding bare states, the real V particle state $|\lambda\rangle$ has to be determined from

$$H|\lambda\rangle = E|\lambda\rangle\,. \tag{3.21}$$

808

Because of N_1, N_2 conservation

$$|\lambda\rangle = \left(-c\psi_V^* + \psi_N^* \int \varphi(\boldsymbol{k})\, a^*(\boldsymbol{k})\, \mathrm{d}\boldsymbol{k}\right)|0\rangle\,. \tag{3.22}$$

Substituting this into (3.21) one finds

$$(m_V - E)c = \frac{g_0}{\sqrt{4\pi}} \int \frac{1}{\sqrt{2\omega}}\, \varphi(\boldsymbol{k})\, \mathrm{d}\boldsymbol{k}\,, \tag{3.23}$$

$$(\omega - E)\varphi(\boldsymbol{k}) = c\, \frac{g_0}{\sqrt{4\pi}}\, \frac{1}{\sqrt{2\omega}}\,. \tag{3.24}$$

If $E = \omega$ we obtain the N — 0 scattering states, which have a positive norm; they can be written down quite easily, but here we do not deal with them. If $E \neq \omega$ from (3.24)

$$\varphi(\boldsymbol{k}) = c\, \frac{g_0}{\sqrt{4\pi}}\, \frac{1}{\sqrt{2\omega}(\omega - E)}\,. \tag{3.25}$$

Substituting this into (3.23) E has to be determined from

$$m_V - E = g_0^2 \int \frac{k^2\, \mathrm{d}k}{2\omega(\omega - E)}\,. \tag{3.26}$$

If we define the function

$$h(z) = \frac{z - m_V}{g_0^2} + \int \frac{k^2\, \mathrm{d}k}{2\omega(\omega - z)}\,, \tag{3.27}$$

the bound V particle energy eigenvalues are given by the zeros of h:

$$h(z) = 0\,, \qquad z \neq \omega\,. \tag{3.28}$$

Using the identity

$$\frac{1}{\omega - z} = \frac{1}{\omega} + \frac{z}{\omega^2} + \frac{z^2}{\omega^2(\omega - z)}\,,$$

one finds

$$\begin{cases} h(z) = a + bz + G(z)z^2\,, \\ a = -\dfrac{m_V}{g_0^2} + \displaystyle\int \frac{k^2\, \mathrm{d}k}{2\omega^2}\,, \\ b = \dfrac{1}{g_0^2} + \displaystyle\int \frac{k^2\, \mathrm{d}k}{2\omega^3}\,, \qquad G(z) = \displaystyle\int \frac{k^2\, \mathrm{d}k}{2\omega^3(\omega - z)}\,, \end{cases} \tag{3.29}$$

$G(z)$ is a finite function. In this model a and b are considered finite parameters, therefore g_0 and m_V behave as ($\hat{\omega}$ is the cut-off energy)

$$g_0^2 \sim -\frac{2}{\log \hat{\omega}}, \qquad m_V \sim -\frac{\hat{\omega}}{\log \hat{\omega}}.$$

For z real and $z < m_1$

$$\frac{d^2k}{dz^2} > 0, \qquad \left.\frac{dh}{dz}\right|_{-\infty} = -\infty, \qquad \left.\frac{dh}{dz}\right|_{m_0} = +\infty, \tag{3.30}$$

thus according to the values of a and b there are three different cases:

1) there are two real roots. (Discrete ghost case);

2) there is only one double root. (Dipole ghost case);

3) there are two complex roots with $E_1 = E_2^*$. (Complex ghost case).

We treat the three cases separately.

1) In the discrete ghost case, one obtains from (3.28) two real eigenvalues E_λ and $E_{-\lambda}$, and from (3.22), (3.25) two eigenvectors $|\lambda\rangle$ and $|-\lambda\rangle$. According to (2.2) they are orthogonal. This can be confirmed also by explicit calculations. From (3.22), (3.25) and (3.27) one obtains

$$\langle \pm\lambda | \pm\lambda \rangle = |c_\pm g_0|^2 \left.\frac{dh}{dz}\right|_{\substack{E_\lambda \\ E_{-\lambda}}}. \tag{3.31}$$

Thus from (3.30) it follows, that choosing $c_{\pm\lambda}$ appropriately

$$\langle \pm\lambda | \pm\lambda \rangle = \pm 1.$$

$|-\lambda\rangle$ is called ghost state.
Since $(dh/dz)|_{E_\lambda}$ is finite

$$|c_\lambda|^2 \sim \log \hat{\omega}.$$

2) In the dipole ghost case there is only one real double root of h to which from (3.22) and (3.25) corresponds only one eigenstate $|\lambda\rangle$. As in (3.31)

$$\langle \lambda | \lambda \rangle = |c_\lambda g_0|^2 \left.\frac{dh}{dz}\right|_{E_\lambda}, \qquad \text{but here} \left.\frac{dh}{dz}\right|_{E_\lambda} = 0.$$

Thus we have found an eigenvector of H with real eigenvalue and zero norm. Because of (2.2) $|\lambda\rangle$ is orthogonal to all other states (scattering eigenstates),

and since the metric is non-degenerate a state $|D\rangle$ having the form (2.8)

$$(H - E_\lambda)|D\rangle = |\lambda\rangle \tag{3.32}$$

necessarily exists. Indeed substituting (3.22) and

$$|D\rangle = \left(-\mathrm{d}\psi_{\mathrm{V}}^* + \psi_{\mathrm{N}}^* \int \Phi(\boldsymbol{k})\, a^*(\boldsymbol{k})\, \mathrm{d}\boldsymbol{k}\right)|0\rangle\,, \tag{3.33}$$

into (3.32) one finds

$$(m_{\mathrm{V}} - E_\lambda)d = \frac{g_0}{\sqrt{4\pi}} \int \frac{1}{\sqrt{2\omega}}\, \Phi(\boldsymbol{k})\, \mathrm{d}\boldsymbol{k} + o_\lambda, \tag{3.34}$$

$$(\omega - E_\lambda)\, \Phi(\boldsymbol{k}) = d\, \frac{g_0}{\sqrt{4\pi}}\, \frac{1}{\sqrt{2\omega}} + \varphi(\boldsymbol{k})\,. \tag{3.35}$$

Using (3.25), (3.35) gives

$$\Phi(\boldsymbol{k}) = \frac{g_0}{\sqrt{4\pi}}\, \frac{1}{\sqrt{2\omega}(\omega - E_\lambda)} \left(d + \frac{o_\lambda}{\omega - E_\lambda}\right). \tag{3.36}$$

Substituting this into (3.34), we obtain that (3.34) is satisfied with an arbitrary d if

$$\left.\frac{dh}{dz}\right|_{E_\lambda} = 0 \tag{3.37}$$

where (3.26) was also taken into account. This shows that a state of the type (3.32) exists only in this case. $|D\rangle$ is called dipole ghost state. From (3.27), (3.33) and (3.36)

$$\langle D|D\rangle = |o_\lambda g_0|^2 \left\{\frac{1}{6}\left.\frac{\mathrm{d}^3 h}{\mathrm{d}z^3}\right|_{E_\lambda} + \frac{1}{2}\left(\frac{d}{o_\lambda} + \frac{d^*}{o_\lambda^*}\right)\left.\frac{\mathrm{d}^2 h}{\mathrm{d}z^2}\right|_{E_\lambda}\right\}.$$

So far d was arbitrary. Choosing it appropriately

$$\langle D|D\rangle = 0 \tag{3.38}$$

can be satisfied. Since the metric is not degenerate, necessarily $\langle\lambda|D\rangle \neq 0$. Indeed

$$\langle\lambda|D\rangle = \frac{|o_\lambda g_0|^2}{2}\left.\frac{\mathrm{d}^2 h}{\mathrm{d}z^2}\right|_{E_\lambda} > 0\,.$$

If

$$|o_\lambda|^2 = \frac{2}{|g_0|^2 (\mathrm{d}^2 h/\mathrm{d}z^2)\,|_{E_\lambda}}, \qquad \langle\lambda|D\rangle = 1. \tag{3,39}$$

Although $|\lambda\rangle$ and $D\rangle$ are both vectors having zero norm, (3.38) and (3.39) seems to be a useful « normalization ». If we suppose that there are other

states of the kind (2.8) necessary to complete the system of energy eigenfunctions

$$(H - E_\lambda)|T\rangle = |D\rangle$$

and substitute a $|T\rangle$ of the form (3.22) or (3.33) we obtain that such a state could exist, only if E were a triple root of $h(z)$. In our case $\mathrm{d}^2h/\mathrm{d}z^2 > 0$ showing that a state $|T\rangle$ does not exist, so the system is now a complete one and the minimum polynomial of H in this subspace is

$$m(H) = (H - E_\lambda)^2 .$$

If $|E'\rangle$ is an energy eigenstate with $E' \neq E_\lambda$, of course $\langle E'|\lambda\rangle = 0$, and from

$$\langle E'|H|D\rangle = E'\langle E'|D\rangle = E_\lambda\langle E'|D\rangle + \langle E'|\lambda\rangle , \qquad \langle E'|D\rangle = 0 .$$

Here (3.32) was used. The projection operator to the subspace $|\lambda\rangle$, $|D\rangle$ is

$$P = \frac{|\lambda\rangle\langle D|}{\langle D|\lambda\rangle} + \frac{|D\rangle\langle\lambda|}{\langle\lambda|D\rangle} , \tag{3.40}$$

and the norm of a state vector $|A\rangle$

$$\langle A|A\rangle = \langle A|P|A\rangle + c , \qquad c \geqslant 0 ,$$

showing that the norm of an arbitrary state vector may be negative only if $|A\rangle$ contains both $|\lambda\rangle$ and $|D\rangle$. From the Schrödinger equation

$$i\frac{\mathrm{d}|t\rangle}{\mathrm{d}t} = H|t\rangle ,$$

it follows

$$|\lambda, t\rangle = |\lambda, 0\rangle \exp[-iE_\lambda t] ,$$

but

$$|D, t\rangle = \{|D, 0\rangle - it|\lambda, 0\rangle\}\exp[-iE_\lambda t] . \tag{3.41}$$

3) In the complex ghost case $h(z)$ has two complex roots E_λ and E_{λ^*}, they are complex conjugates: $E_\lambda = E_{\lambda^*}^*$. to which, because of (3.22) and (3.25), correspond two eigenstates $|\lambda\rangle$ and $|\lambda^*\rangle$. Because of (2.1)

$$\langle\lambda|\lambda\rangle = \langle\lambda^*|\lambda^*\rangle = 0 .$$

Indeed for example

$$\langle \lambda | \lambda \rangle = \frac{|c_\lambda g_0|^2}{\operatorname{Im} E_\lambda} \operatorname{Im} h(z)|_{E_\lambda} = 0 .$$

Since the metric is non degenerate, necessarily $\langle \lambda^* | \lambda \rangle \neq 0$.

$$\langle \lambda^* | \lambda \rangle = c^*_{\lambda^*} c_\lambda |g_0|^2 \frac{\mathrm{d}h}{\mathrm{d}z}\Big|_{E_\lambda} \neq 0 .$$

The projection operator has also here the form

$$P = \frac{|\lambda\rangle\langle\lambda^*|}{\langle\lambda^*|\lambda\rangle} + \frac{|\lambda^*\rangle\langle\lambda|}{\langle\lambda|\lambda^*\rangle}, \tag{3.42}$$

and the norm of an arbitrary vector may be negative only if it contains both $|\lambda\rangle$ and $|\lambda^*\rangle$. Similarly

$$\langle A | H | A \rangle = E_\lambda \frac{\langle A|\lambda\rangle\langle\lambda^*|A\rangle}{\langle\lambda^*|\lambda\rangle} + \text{c.c.} + \ldots .$$

Thus one is inclined to say that the probability of measuring the complex E_λ is the (complex)

$$W_\lambda = \frac{\langle A|\lambda\rangle\langle\lambda^*|A\rangle}{\langle\lambda^*|\lambda\rangle},$$

but it is different from zero only if $|A\rangle$ contains both eigenstates. Then the two members together give a real value.

3·7. *Other field theories.* – Beside the Lee model there are other model theories, which require also a quantization with indefinite metric. The models in question are: the model of MACHIDA [28], the model of GOLDSTEIN [29], the RUIJGROOK-VAN HOVE model [30] and the model of TEODOROVICH [31]. These models tend, essentially, to extend the Lee model towards more realistic field theories. Here we do not wish to describe these models in detail, only we mention that *e.g.* [29] extends the Lee model by allowing the meson to create one nucleon pair. This results in a shifting of the ghost state from the V particle spectrum to the meson spectrum. Therefore the ghost was nicknamed as « galloping ghost ». In [31] on the other hand the Lee model was generalized by adding to the Hamiltonian (3.19) a term $E_\nu \int \varphi^* \varphi \mathrm{d}p - g_\nu \int (\varphi^* \psi_V - \psi^* \varphi^*_V)\, \mathrm{d}p$ were φ is the field operator of a new type of particle, having a usual anticommutator. Choosing E_α and g_α appropriately also a « tripole » ghost state can be found [32].

Other model theories, however, do not require indefinite metric (infinitely heavy nucleon interacting with scalar mesons, the LEE model with a local interaction [33], the THIRRING model [34], WENTZEL's pair theory [35]. The question naturally arises whether the realistic field theories require an indefinite metric or not. According to some authors [36] a contradiction of the type (3.18) occurs also in real field theories. The questions of ghost states in these theories were analyzed by KÄLLEN [37] and by FORD [38], nevertheless these problems are not yet completely answered.

3·8. *Heisenberg's theory.* – In the HEISENBERG theory of elementary particles [17], there are some indications that the field equation should be completed by supposing a form of the type

$$\frac{1}{2} S'(p) = \int \varrho(\varkappa^2)\, \mathrm{d}\varkappa^2 \left| \frac{\gamma_\mu p_\mu + i\varkappa}{p^2 + \varkappa^2} - \frac{\gamma_\mu p_\mu + i\varkappa}{p^2} + \frac{\gamma_l p_\mu \varkappa^2}{(p^2)^2} \right],$$

for the Fourier transform of the vacuum expectation value:

$$S'(x, x') = \langle 0 | \{ \psi(x), \overline{\psi}(x') \} | 0 \rangle .$$

This result should be calculated, at least in principle, from the theory itself.

The form of S' requires obviously a quantization by means of an indefinite metric, since according to LEHMAN and KÄLLEN [39] S' has otherways the form

$$S' \sim \int \varrho(\varkappa^2)\, \mathrm{d}\varkappa^2\, S(\varkappa^2) \qquad \text{with} \qquad \varrho(\varkappa^2) \geqslant 0 .$$

3·9. *Markov's proposal.* – In two consecutive papers [21-22], MARKOV considers field equations, where instead of the usual quadratic form $s^2 = -x_\mu^2$ the modified $s'^2 = -(x_\mu^2 + a^2)$ stands, where a^2, an invariant, characterizes the mean extension of a field particle, which may be different, and characteristic for different particle families. Since the substitution

$$x'_\mu = x_\mu \frac{s'}{s},$$

just produces s'^2, for the simplest massless scalar boson particle the « extended » particle free field equation was proposed to have the form

$$\Box \varphi = 0$$

and quite similarly for more complicated particles. This leads to the conse-

quence that the Green functions of this theory are the ordinary ones, but everywhere s'^2 stands instead of s^2. Thus, for example,

$$D \to D\ (s'^2) = \frac{-\varepsilon(s'^2)}{2\pi}\,\delta(s'^2)\,,$$

$$D^1 \to D^1(s'^2) = \frac{1}{(2\pi)^3}\int \delta(k_\mu^2)\,\exp\,[ikx']\,\mathrm{d}k = \frac{-1}{(2\pi)^2}\int_0^\infty \sin\,(s'^2\alpha)\,\mathrm{d}\alpha\,,$$

$$\bar{D} \to \bar{D}\ (s'^2) = \frac{1}{(2\pi)^4} P\int \frac{\exp\,[ikx']}{k_\mu^2}\,\mathrm{d}k = \frac{1}{4\pi^2}\int_0^\infty \cos\,(s'^2\alpha)\,\mathrm{d}\alpha\,,$$

$$D_c \to D_c(s'^2) = \bar{D} + \frac{1}{2}\,iD^1 = \frac{1}{4\pi^2}\int_0^\infty \exp\,[i\alpha s'^2]\,\mathrm{d}\alpha\,.$$

From these $D^1(k)$ defined by

$$D^1(k) = \frac{1}{(2\pi)^4}\int D^1(x)\,\exp\,[ikx]\,\mathrm{d}x$$

is

$$D^1(k) = \frac{2}{(2\pi)^4}\int \cos\left(k_\mu^2\,\beta - \frac{a^2}{4\beta}\right)\mathrm{d}\beta\,.$$

But, on the other hand,

$$D^1(x) = \int D^1(k)\,\exp\,[ikx]\,\mathrm{d}k = \iint \mathrm{d}\varkappa^2\,\delta(k^2+\varkappa^2)\,D^1(-\varkappa^2)\,\exp\,[ikx]\,\mathrm{d}k =$$

$$= \int \mathrm{d}\varkappa^2\,\varrho(\varkappa^2)\,\Delta^1(\varkappa^2)\,,$$

where

$$\varrho(\varkappa^2) = (2\pi)^3 D^1(-\varkappa^2) = \frac{1}{\pi}\int_0^\infty \cos\left(-\varkappa^2\beta - \frac{a^2}{4\beta}\right)\mathrm{d}\beta = \delta(\varkappa^2) - \frac{1}{2}\,a^2\,\mathrm{Re}\,\frac{H_1^{(1)}(a\varkappa)}{a\varkappa}\,.$$

Since $\varrho(\varkappa^2)$ may also be negative, referring to the Lehman-Källen theorem, Markov stresses that this is a theory with indefinite metric.

This is, however, not at all certain. In this theory one possesses an energy-momentum four-vector with the property

$$[P_\mu, \varphi] = i\,\partial'_\mu \varphi$$

which also violates the conditions of the Lehman-Källen theorem and makes everything understandable without using an indefinite metric.

Whether an indefinite metric is unavoidable here or not would be quite clear only ofter treating interaction problems; these questions are, however not yet worked out in detail. Our opinion is that in this theory it is not necessary to introduce an indefinite metric. We will not deal with this proposal in the following since without the inclusion of an interaction one cannot yet form a definite judgement.

3·10. *Froissart's model.* – In this part we treat a simple relativistic field theoretical example given by FROISSART [23]. This model serves for the better understanding of the « pole-type » ghost states, one of which, with the dipole ghost, we have met already in the Lee model. This model is of a quite different type, but relativistically invariant.

Let us start with the Lagrangian

$$L = -(\partial_\mu A \partial_\mu B + \mu^2 AB + \lambda^2 A^2)$$

where A and B are two scalar fields. From this Lagrangian, following the usual way, the field equations and the equal time commutators are given

$$(\Box - \mu^2)A = 0\,, \qquad (\Box - \mu^2)B = 2\lambda^2 A\,,$$

$$[A, A] = [B, B] = [\dot{A}, A] = [A, B] = [\dot{B}, B] = 0\,,$$

$$[\dot{A}(\boldsymbol{x}, t), B(\boldsymbol{x}', t)] = [\dot{B}(\boldsymbol{x}, t), A(\boldsymbol{x}', t)] = -i\,\delta(\boldsymbol{x} - \boldsymbol{x}')\,.$$

The solutions of the field equations are

$$A(x) = A_0(x) = \sum \sqrt{\frac{1}{2\omega}}\,(a(k)\exp[ikx] + a^*(k)\exp[-ikx])\,, \qquad k_\mu^2 = -\mu^2,$$

$$B(x) = B_0(x) + \frac{\lambda^2}{\mu^2}(1 + x_\mu \partial_\mu) A_0(x)\,,$$

$$B_0(x) = \sum \sqrt{\frac{1}{2\omega}}\,(b(k)\exp[ikx] + b^*(k)\exp[-ikx])\,, \qquad k_\mu^2 = -\mu^2.$$

The equal time commutator are satisfied if

(3.42′) $\qquad [a, a] = [a, a^*] = [b, b] = [b, b^*] = [a, b] = 0\,, \qquad [a(k), b^*(k')] = \delta_{kk'}.$

From these

$$[A(x), A(x')] = 0\,, \qquad [A(x), B(x')] = i\,\Delta(x - x')\,,$$

$$[B(x), B(x')] = i\,\frac{\lambda^2}{\mu^2}[2 + (x_\mu - x'_\mu)\partial_\mu]\,\Delta(x - x') = 2i\lambda\,\frac{\partial}{\partial\mu^2}\,\Delta(x - x')\,.$$

An energy-momentum operator can be defined by

$$[P_\mu, A] = i\partial_\mu A\,, \qquad [P_\mu, B] = i\partial_\mu B\,,$$

which in momentum representation is:

$$P_\mu = \sum k_\mu \left\{a^*(k)b(k) + b^*(k)a(k) + \frac{\lambda^2}{\mu^2}\, a^*(k)a(k)\right\}. \tag{3.43}$$

From (3.43) it follows

$$[P_\mu, B_0] = i\,\partial_\mu \left\{B_0 + \frac{\lambda^2}{\mu^2} A_0\right\}. \tag{3.44}$$

Now we define the vacuum state $|0\rangle$ for which $P_\mu|0\rangle = 0$. This is satisfied if as usual

$$a|0\rangle = b|0\rangle = 0\,, \qquad \langle 0|0\rangle = 1\,.$$

α. The one particle sector.
There are two one particle states

$$|1, 0\rangle = a^*(k)|0\rangle\,, \qquad |0, 1\rangle = b^*(k)|0\rangle\,.$$

From (3.43) or (3.44) it follows

$$(P_\mu - k_\mu)|1, 0\rangle = 0\,, \qquad (P_\mu - k_\mu)|0, 1\rangle = \frac{\lambda^2}{\mu^2}\, k_\mu |1, 0\rangle\,, \tag{3.45}$$

and from (3.42)

$$\langle 1, 0|1, 0\rangle = \langle 0, 1|0, 1\rangle = 0\,, \qquad \langle 1, 0|0, 1\rangle = \delta_{kk'}\,, \tag{3.46}$$

showing the typical dipole-ghost character of $|0, 1\rangle$. In this one particle sector, the « minimum polynomial » is

$$(P_\mu - k_\mu)^2\,.$$

β. The two particle sector.
The two particle eigenstates are

$$|2, 0\rangle = a^*(k_1)a^*(k_2)|0\rangle\,, \qquad |0, 2\rangle = b^*(k_1)b^*(k_2)|0\rangle\,,$$
$$|1, 1\rangle = a^*(k_1)b^*(k_2)|0\rangle\,.$$

815

From (3.43) these satisfy

$$\{P_\mu - (k^1_\mu + k^2_\mu)\}|2, 0\rangle = 0\,,$$

$$\{P_\mu - (k^1_\mu + k^2_\mu)\}|1, 1\rangle = \frac{\lambda^2}{\mu^2}\, k^2_\mu |2, 0\rangle,$$

$$\{P_\mu - (k^1_\mu + k^2_\mu)\}|0, 2\rangle - \frac{\lambda^2}{\mu^2}\{k^1_\mu a^*(k_1) b^*(k_2) + k^2_\mu b^*(k_1) a^*(k_2)\}|0\rangle\,.$$

Thus $|1, 1\rangle$ may be called dipole; $|0, 2\rangle$ tripole ghost. From (3.42)

$$\text{(3.47)} \qquad \begin{cases} \langle 2, 0|2, 0\rangle = \langle 2, 0|1, 1\rangle = \langle 0, 2|0, 2\rangle = 0\,, \\ \langle 1, 1|1, 1\rangle = \delta_{k_1 k_4}\delta_{k_2 k_3}\,, \\ \langle 2, 0|0, 2\rangle = \delta_{k_1 k_4}\delta_{k_2 k_3} + \delta_{k_1 k_3}\delta_{k_2 k_4}\,. \end{cases}$$

In this subspace « the minimum polynomial » is

$$\{P_\mu - (k^1_\mu + k^2_\mu)\}^3\,.$$

These results are in agreement, of course, with the general principles. Indeed let us suppose that there exist three state vectors, defined by

$$(H - E)|1\rangle = 0\,,$$

$$(H - E)|2\rangle = |1\rangle\,,$$

$$(H - E)|3\rangle = |2\rangle\,.$$

In this case $\langle 1|1\rangle = 0$, since just this was the necessary condition of the system of eingenfunctions being incomplete. From the third equation $\langle 1|(H - E)|3\rangle = \langle 1|2\rangle$, from the first $\langle 3|(H - E)|1\rangle = 0$. Thus $\langle 2|1\rangle = 0$. But in this case $\langle 1|3\rangle \neq 0$, since the metric is non degenerate. Now from the third equation $\langle 2|(H - E)|3\rangle = \langle 2|2\rangle$, from the second $\langle 3|(H - E)|2\rangle = \langle 3|1\rangle$. Thus $\langle 2|2\rangle = \langle 1|3\rangle$ = real number $\neq 0$. From the third equation $\langle 3|2$ is real and may be $= 0$, and also $\langle 3|3\rangle = 0$ without violating the requirement of the non-degenerateness of the metric.

The higher sectors of the above model can be treated similarly; there newer and newer ghost states emerge, with higher and higher « pole order ».

4. – Ghost contributions to propagators.

Previously we have seen field theoretical examples for field theories with an indefinite metric. There were several examples among them in which we have referred to the Källen-Lehman spectral representation of the propa-

gators. Since, besides, in field theories the propagators play a central rôle, we study here the gost state contributions appearing in the spectral representation of the propagators.

4·1. *Contributions from ghosts of known types.* – In this part we treat the different possible contributions to the propagators arising from ghost states which we have discussed earlier. For the sake of simplicity we treat the situation in the case of a relativistic hermitian scalar (pseudoscalar) field φ; the results can be, however, quite easily generalized to other cases. For the calculations the Källen-Lehman technique is applied, in the spirit of which the Heisenberg picture is used throughout.

We suppose that there exists an energy-momentum four vector with the properties:

$$[P_\mu, \varphi] = i\partial_\mu \varphi\,, \qquad [P_\mu, P_\nu] = 0\,, \qquad P_i^* = P_i\,, \qquad P_0^* = P_0\,, \tag{4.1}$$

and that therefore the P_μ can be diagonalized simultaneously and we assume a non-degenerate metric.

Now, according to the general discussions given in part 2, there are two possibilities:

(i) the system of the eigenfunctions of P_μ is complete;

(ii) the system is not complete.

We treat case (i) first. In this case there are also two other possibilities:

(a) all the eigenvalues of P_i, P_0 are real;

(b) some of the eigenvalues are complex.

Let us pick up one eigenvalue p with $p^2 = -M^2$ where $-M^2$ is the square of the four eigenvalue, and suppose $p_0 > 0$ because the eigenvalues of the energies are positive definite (cfr. (3.7′)), and let us treat the case (a) first.

Since the metric is not degenerate, one can always orthonormalize the eigenvectors corresponding to P_μ to ± 1. Indeed, if there were one eigenvector $|p\rangle^0$ with zero norm, assuming a non-degenerated metric, one has to find at least one other vector $|p\rangle'$ which is not orthogonal to $|p\rangle^0$. According to (2.2) $|p\rangle'$ has to correspond to the same eigenvalue p_μ: From the combination of these state vectors, one can always build up two orthogonal eigenvectors with norms $+1$ or -1. The orthonormalization procedure can never break up, for we have assumed the metric to be non-degenerate. Considering now that in this case the unit operator

$$1 = \sum |p, \alpha\rangle(\pm 1)\langle p, \alpha|$$

where α denotes the other indices necessary to specify the state, one obtains in the usual way

$$\Delta' = \langle 0 | [\varphi(x), \varphi(x')] | 0 \rangle = \int \varrho(\varkappa^2)\, \Delta(x - x'; \varkappa^2)\, \mathrm{d}\varkappa^2 ,$$

but

$$\varrho(\varkappa^2) \gtreqless 0 .$$

If there exists one stable particle state $p^2 = -\mu^2$ with the norm ± 1, of course $\varrho = \pm\, \delta(K^2 - \mu^2) + \sigma(\varkappa^2)$.

Consider now the case (*b*) and a fixed complex eigenvalue p. The corresponding eigenvector $|p\rangle$, because of (2.1) has zero norm. Assuming non-degenerate metric at least one state vector non orthogonal to $|p\rangle$ necessarily exists. This, considering (2.1), corresponds to the eigenvalue p^*. Thus, if the metric is non-degenerate, both p and p^* are eigenvalues. The projection operator to this subspace has the form (cfr. (3.42))

$$P \sim \frac{|p\rangle\langle p^*|}{\langle p^* | p\rangle} + \frac{|p^*\rangle\langle p|}{\langle p | p^*\rangle} ,$$

where $P_\mu |p^*\rangle = p^*_\mu |p^*\rangle$ and since p is complex, Δ' is of a « decaying » type.

In the case of (ii) we may find also the two possibilities (*a*) and (*b*). Considering only real eigenvalues [40] because of (2.7) an eigenvector with zero norm necessarily exists.

$$(4.2) \qquad (P_\mu - p_\mu)|p\rangle = 0 , \qquad \langle p | p\rangle = 0 , \qquad p_\mu^2 = -\mu^2 , \qquad p_0 > 0 .$$

For completing the vector space one has to add a state of type (2.8) (dipole ghost state) defined by

$$(4.3) \qquad (P_\mu - p_\mu)|pD\rangle = c_\mu |p\rangle , \qquad \langle pD | pD\rangle = 0 ,$$

where because of relativistic invariance

$$c_\mu = f(p^2) p_\mu .$$

This is obviously the relativistic generalization of the definition of a dipole ghost state of the type (2.8) (cfr. eq. (3.45)). Since the metric is non-degenerate

$$\langle pD | p'\rangle = a \delta_{pp'} , \qquad a \neq 0 ,$$

but, since both vectors have zero norm, the values f and a are arbitrary. Nevertheless, since $\langle pD | P | pD\rangle$ must be real, fa is also real. In this case

the projection operator to this subspace is

$$P = \sum \left\{ \frac{|\boldsymbol{p}D\rangle\langle \boldsymbol{p}|}{a} + \frac{|\boldsymbol{p}\rangle\langle \boldsymbol{p}D|}{a^*} \right\},$$

and therefore the contribution of these states to Δ' is

$$\Delta' \sim \sum \left\{ \langle 0|\varphi(x)|\boldsymbol{p}D\rangle \frac{1}{a} \langle \boldsymbol{p}|\varphi(x')|0\rangle + \langle 0|\varphi(x)|\boldsymbol{p}\rangle \frac{1}{a^*} \langle \boldsymbol{p}D|\varphi(x')|0\rangle \right\} - \{x \rightleftarrows x'\} .$$

From (4.1) and (4.2) one obtains as usual

$$\langle 0|\varphi(x)|p\rangle = \langle 0|\varphi(0)|p\rangle \exp[ipx] ,$$

but from (4.1) and (4.3)

$$\langle 0|\varphi(x)|pD\rangle = \{\langle 0|\varphi(0)|pD\rangle - ic_\mu x_\mu \langle 0|\varphi(0)|p\rangle\} \exp[ipx]$$

showing the peculiar x dependence found already in the Lee model (eq. (3.41)) and also in Heisenberg's theory. Substituting the matrix elements into the expression for Δ' and requiring relativistic invariance, it follows

$$\Delta'(x - x') \sim C_1 \Delta(x - x') + C_2 (x - x')_\mu \partial_\mu \Delta(x - x') ,$$

C_1, C_2 are real, C_1 may also be negative, sign C_2 = sign fa. Taking the Fourier transform, the form of the contribution to the states of the type (4.2) and (4.3) is

$$\Delta'(p)) \sim \int \left\{ \varrho_1(\varkappa^2) \frac{1}{p^2 + \varkappa^2} + \varrho_2(\varkappa^2) \frac{\varkappa^2}{(p^2 + \varkappa^2)^2} \right\} \mathrm{d}\varkappa^2 , \tag{4.4}$$

both functions ϱ_1 and ϱ_2 may be negative.

In the case of a non-complete system of eigenfunctions with complex eigenvalues, the properties of the two previous cases appear together.

Contributions from ghost states with higher pole order can also be calculated in a similar fashion [32].

4·2. *Contributions found by certain approximation methods.* – As it was already frequently mentioned, if there are only normal states, according to the Källen-Lehman's results the meson propagator, for example, has the general form

$$\Delta'_F = \frac{1}{p^2 + \varkappa^2} + \int \frac{\sigma(\varkappa^2)\, \mathrm{d}\varkappa^2}{p^2 + \varkappa^2} .$$

In this equation $\sigma(\varkappa^2)$ is a real and positive function for $\varkappa_0^2 < \varkappa^2 < \infty$ where $\varkappa_0^2 > 0$. The propagator, therefore, has a pole at $p^2 = -\varkappa^2$ and a branch-line going from $\varkappa_0^2$ to $+\infty$. Using a certain approximation method, considering only those diagrams for the meson propagator, which can be represented by inserting an arbitrary number of second order nucleon loops, an expression can be obtained for Δ_F' which has an additional pole. It was pointed out that Δ_F' obtained in this way can be rewritten as a well behaved (in the sense of Källen-Lehman theorem) part, and an additional contribution. This was called ghost contribution [41-43]. The reason of the appearance of this additional contribution (since, of course, an exact solution is not known) may be:

(i) the exact consequence of the field theory;

(ii) the exact solution is not defined in a conventional field theory thus one may add additional postulates eliminating the unnecessary contributions.

(iii) the result of a bad approximation procedure.

The cases (ii) and (iii) are out of our interest, and by an appropriate procedure one may eliminate the ghost contribution. In the case (i) one possesses presumably a theory with indefinite metric as we have already mentioned in Sect. 3·7. In this case, if the unpleasant contribution is eliminated, the propagator thus obtained corresponds to a field-theory which is different from the one with which the calculation was started.

It was shown by MEDVEDEV and POLIVANOV [44] in the case of the Lee model discrete ghost case, that by omitting the ghost contribution from the propagator, the resulting new propagator corresponds to a Hamiltonian with a cut-off function $f(\omega)$ for which the contradiction (3.18) does not occur, *i.e.*

$$|\sigma| = \frac{g^2}{g_0^2} = 1 - g^2 \int \frac{f(\omega) k^2 \, \mathrm{d}k}{2\omega(E - \omega)} > 0 \, .$$

5. – Difficulties connected with the probabilistic interpretation.

In this part we discuss the difficulties, which according to Sect. 2·3 arise or may arise in a theory with an indefinite metric. Proposals for eliminating these difficulties will also be treated.

5·1. *Decomposition of the state vector space.* – We have seen that state vectors with positive or zero norm lead to difficulties. Therefore one has to consider some decompositions of the state vector space into « physical » and « non-

physical » subspaces. Let the vector space be spanned by $|i\rangle$, ... Then

$$1 = \sum |i+\rangle\langle i+| - \sum |i-\rangle\langle i-|$$

gives a decomposition to two subspaces $\mathscr{H}_+$ and $\mathscr{H}_-$. $\mathscr{H}_+$ spanned by $|i+\rangle$, ... is a Hilbert space, $\mathscr{H}_-$ a vector space with negative metric. The decomposition is, however, obviously not unique [45-46]. Indeed, let, for example,

$$\left.\begin{aligned} |a+\rangle &= \sqrt{2}|\lambda+\rangle + |1-\rangle, \quad \langle a+|a+\rangle = 1, \\ |a-\rangle &= \sqrt{2}|1-\rangle + |1+\rangle, \quad \langle a-|a-\rangle = -1, \end{aligned}\right\} \langle a+|a-\rangle = 0,$$

then $|a+\rangle$, $|2+\rangle$, ... and $|a-\rangle$, $|2-\rangle$, ... give another decomposition. Therefore we need some other information to obtain a definite decomposition. But we want to describe the real world, or in model-theories, we wish to maintain the maximum number of the original features of the model. The decompositions thus obtained are called natural decompositions. The subspace containing only physical states will be denoted by $\mathscr{H}_p$ (in Heisenberg's work H_I) the complementary subspace $\mathscr{H} - \mathscr{H}_p$ by $\mathscr{H}_n$, with corresponding projection operators P_p and P_n, respectively. Then

$$|\rangle = |\rangle_p + |\rangle_n = P_p|\rangle + P_n|\rangle .$$

Now we treat the decomposition in the examples discussed in Sect. 3.

3·3. *Electromagnetic field.* – Since longitudinal and scalar photons are unobservable, the state vectors

$$|N_1, N_2, 0, 0\rangle \text{ span the space } \mathscr{H}_p,$$

$$|N_1, N_2, N_3, N_0\rangle,\ N_3 \neq 0,\ N_0 \neq 0 \text{ or both, span } \mathscr{H}_n .$$

A state satisfying the Lorentz condition is called an element of the Lorentz set. For such a state

$$|L\rangle = |L\rangle_p + |L\rangle_n, \qquad {}_n\langle L|L\rangle_n = {}_n\langle L|L\rangle_p = 0 . \tag{5.1}$$

3·4. *Higher order theory.* – Since $\varphi^{(1)}$ is a regularizing auxiliary field, its particles are not observable. Therefore

$$|N_0, 0\rangle \text{ span the space } \mathscr{H}_p,$$

$$|N_0, N_1\rangle,\ N_1 \neq 0 \text{ span } \mathscr{H}_n .$$

3·5. *Auxiliary field.* – Here the decomposition is obviously the same as in the previous case.

3·6. *The Lee model.* – This is a model theory in which different natural decompositions may exist. First, of course, it is possible that the bare particle states

$$|0, N_N, N_\theta\rangle \text{ span the space } \mathscr{H}_p$$

$$|N_V, N_N, N_\theta\rangle,\ N_V \neq 0 \text{ span } \mathscr{H}_n$$

similarly as before. Later, when we treat the possibilities of the probabilistic interpretation we treat another even more important decomposition.

3·8. *Heisenberg's theory.* – States giving the contribution of the type

$$\frac{\gamma_\mu p_\mu \varkappa^2}{(p^2)^2} - \frac{\gamma_\mu p_\mu + i\varkappa}{p^2}, \tag{5.2}$$

are elements of $\mathscr{H}_n$: Those, resulting normal contributions, span the space $\mathscr{H}_p$.

3·10. *Froissart's model.* – The energy-momentum four-vector P_μ is measurable. Therefore its eigenvectors

$$|N, 0\rangle \text{ span the space } \mathscr{H}_p,$$

$$|N, M\rangle,\ M \neq 0 \text{ span } \mathscr{H}_n.$$

5·2. *Probabilistically interpretable and non-interpretable theories.* – So far we have discussed the decomposition of the state vector space into physical and non-physical subspaces. Now a theory is probabilistically interpretable [46-48] if it fulfils:

(i) a real physical state of a system considered, say, at the time t_0 is described by a state vector $|t_0\rangle_p \in \mathscr{H}_p$, and

(ii) if $|t_0\rangle = |t_0\rangle_p$ then $|t\rangle = |t\rangle_p + |t\rangle_n$ for each t so that ${}_n\langle t|t\rangle_n = {}_n\langle t|t\rangle_p = 0$.

Since the energy is an observable, it is represented by a hermitian operator $H = H^*$. Therefore, according to the Schrödinger equation, the norm is conserved

$$i\frac{\mathrm{d}|t\rangle}{\mathrm{d}t} = H|t\rangle, \qquad \langle t|t\rangle = \text{const}.$$

But the physical part of the state vector is only a distinguished component of the whole vector, therefore it satisfies another equation, namely (taking expediently the interaction picture)

$$i\frac{\mathrm{d}|t\rangle_p}{\mathrm{d}t} = (P_p H(t) P_p + W(t))|t\rangle_p , \tag{5.3}$$

if $\mathrm{d}P/\mathrm{d}t = 0$ and $|-\infty\rangle_n = 0$, where

$$W(t) = P_p H(t) P_n U(t) P_p (P_p U(t) P_p)^{-1} .$$

$H(t)$ is the interaction Hamiltonian, $U(t)$ is the evolution operator $|t\rangle = U(t)|-\infty\rangle$, $U^*(t)\,U(t) = 1$, which is supposed to exist at least formally for the whole theory. $(P_p U(t) P_p)^{-1}$ is defined by

$$(P_p U P_p)^{-1} P_p U P_p = P_p U P_p (P_p U P_p)^{-1} = P_p .$$

$P_p H(t) P_p$ is hermitian, $W(t)$ in general is not. This makes the norm of $|t\rangle_p$ change in time. (The situation is obviously the same also in a vector space with definite metric; cfr. *e.g.* the complex potential of a nucleus considering only one reaction channel.) The solution of (5.3) is, of course,

$$|t\rangle_p = P_p U(t) P_p |-\infty\rangle_p .$$

In an S-matrix theory the postulates (i) and (ii) have to be fulfilled only for $t_0 = -\infty$ and $t = +\infty$, therefore there is no problem with negative scattering probability in a theory, where

$$_p\langle |P_p S^* P_p S P_p| \rangle_p = 1 , \quad \text{provided} \quad {}_p\langle | \rangle_p = 1 \tag{5.4}$$

$|\rangle_p$ is arbitrary in $\mathscr{H}_p$.

We now show that quantum electrodynamics, Froissart's model, are probabilistically interpretable, and discuss the Lee model dipole ghost case and Heisenberg's theory.

It is not difficult to prove [49] that in quantum electrodynamics if $|-\infty\rangle$ is an element of the Lorentz set, $|+\infty\rangle$ becomes also an element. Therefore considering (5.1), (5.4) is fulfilled.

The Froissart's model is interpretable, but obviously trivial. Considering the decomposition in the last paragraph and that $A(x)$ satisfies a free field equation, the states $|N, 0\rangle$ are stationary.

In reference [23] some other speculations are also given on what happens if the fields $A(x)$ and $B(x)$ are certain asymptotic (in- or outgoing) fields in a theory with an interaction. Considering the four momentum conservation

theorem, which the particles of the field B do not fulfil, it seems that the probabilistic interpretation here also can be maintained.

The Lee model dipole ghost case is interpretable only if in the higher sectors bound states with negative norm do not appear. This can be seen most easily by considering in the initial and final scattering states not overlapping wave packets of real particles, which, on the other hand, are broad enough to be considered asymptotically as eigenstates of the total energy H with additive eigenvalues. If now in the initial state there is no dipole ghost wave packet (*i.e.* $|-\infty\rangle \in \mathscr{H}_p$), the energy conservation forbids its appearing also in the final state; the dipole ghost being not an eigenstate of H. Thus, according to (3.40), negative probabilities do not occur. The same is true also for the modification of the Lee model [31], but only for the tripole ghost case [32]. If, however, there are bound states with negative norm in the higher sectors, the above reasoning fails. Such bound states were indicated in the case of the « deuteron » problem, *i.e.* in the sector $\{VN, NN\theta\}$ by VAKS [50].

In Heisenberg's theory we can apply the same reasoning. States giving the contribution (5.2) cannot be eigenfunctions of the energy momentum four-vector, since if they were they would give a contribution of a normal type. So they do not appear in the end state, if initially the state was a normal one. Here, however, one has to be careful. Heisenberg's propagator is similar, but not identical to the propagator of known types treated in Sect. 4·1. Therefore some of the general principles proved to be valid so far might be violated. Anyway, if in this theory a propagator of the type (4.4) were accepted, the theory would be probabilistically interpretable. Indeed, recently, MITTER [51] in a certain approximation has obtained a propagator in this theory — different from (5.2) — which really corresponds to a dipole ghost.

The other theories treated in detail in Sect. 3 are not probabilistically interpretable. For most of them this is quite obvious; we treat here only the Lee model complex ghost case (cfr. reference [16]) using similar simple considerations as in the dipole ghost case.

Consider the scattering 3θ on $2N$. The initial energy eigenvalue is then real. In the end state then beside the state $|2N, 3\theta\rangle$ the wave packets $|\lambda\rangle$ around the point let us say x_1, $|\lambda^*\rangle$ around the point x_2 + one θ packet may appear and the converse also happens. Since the eigenvalues E_λ and E_{λ^*} are complex conjugates, thus the whole energy is real. The superposition of these states $C_1|\lambda(x_1), \lambda^*(x_2), \theta\rangle + C_2|\lambda(x_2), \lambda^*(x_1), \theta\rangle$ both having zero norm but being not orthogonal may give any norm different from zero, positive or negative. The situation in the discrete ghost case is of course even more simple owing to the fact that there the ghost state with negative norm has a real energy eigenvalue.

5·2. *Proposals for eliminating the difficulties.* – We have seen that among the different field theoretical examples — and with the exception of quantum electrodynamics — only the « pole ghost » type theories are probabilistically interpretable. Since theory and experiments in a quantum theory are connected just by the probabilistic interpretation, therefore either we give up the other type of theories as unphysical or try to modify them in order to obtain an interpretable theory. In the following several modifications will be mentioned.

In a non interpretable theory (5.4) was not fulfilled. Therefore if we want to obtain a probabilistically interpretable theory also in this case, we have to construct in a reasonable way a « physical » $\tilde{S}$-matrix, different from $P_p S P_p$, with the property (5.4).

A sufficient condition to the filfulment of (5.4) is, of course:

$$\tilde{S}^* \tilde{S} = P_p \,. \tag{5.5}$$

For the special case of the Lee model discrete ghost case, such an $\tilde{S}$ (in a non-unique manner) was proposed by FERRETTI [52]. In the following we treat in detail only the prescription of BOGOLJUBOV *et al.* [18-20], which can be applied to each theory. The decomposition of the state vector space in this case is just the same as that given in Sect. 5·1.

In quantum electrodynamics the non physical part of a state vector may be quite arbitrary, provided it is an element of the Lorentz set. In the general case of a non interpretable theory, Bogoljubov assumes in contrast to this that the non physical part $|\rangle_n = P_n|\rangle$ of a state vector at $t = -\infty$ is determined by the physical part by means of:

$$|-\infty\rangle_n = N|-\infty\rangle_p \,, \qquad P_n N P_p = N \,, \tag{5.6}$$

with an appropriate N. In this case

$$|+\infty\rangle_p = P_p|+\infty\rangle = P_p S(P_p + N)|-\infty\rangle_p \,, \qquad i.e. \tilde{S} = P_p S(P_p + N) \,. \tag{5.7}$$

N has to be determined from (5.5). This, however, does not determine N uniquely. Because of this arbitrariness, it was shown [12, 53] that certain types of theories possessing a higher order Lagrangian, if these prescriptions are applied, are equivalent to non local theories in the sense that they give the same scattering probabilities, if N is appropriately chosen. Bogoljubov proposes the following procedure for determining N. The norm of the whole state vector according to Sect. 5·2 is, of course, conserved. The norm of the physical part alone also will be conserved, if $|-\infty\rangle_n + |+\infty\rangle_n = 0$. From

the definition of the physical and non-physical parts and from (5.6) it follows

$$0 = |+\infty\rangle_n + N\,|-\infty\rangle_p = P_n S\,|-\infty\rangle + N\,|-\infty\rangle_p = \\ = \{P_n S(P_p + N) + N\}\,|-\infty\rangle_p ,$$

from which

$$N = -(P_n S P_n + P_n)^{-1}\, P_n S P_p , \tag{5.8}$$

and the physical $\tilde{S}$-matrix

$$\tilde{S} = P_p S[P_p - (P_n + P_n\, S P_n)^{-1}\, P_n S P_p] . \tag{5.9}$$

MEDVEDEV and POLIVANOV have shown [20], using the field theoretical model described in Sect. 3·5, that from (5.9) finite results can be obtained.

According to (5.6) and (5.8) the preparation of the initial state in this proposal depends on the process to be studied. Therefore MAKSIMOV has proposed [54] another procedure for obtaining another physical $\tilde{S}$-matrix in form $\tilde{S} = USU^*$, where the operators U and U^* serve to project the state vector before and after the scattering (led by S) so that the norm of the physical part should be conserved. He has not succeeded in solving the equation for U obtained from this requirement in general, therefore this method was applied only to the Lee model, and, essentially, to the example treated in Sect. 3·4.

When the physical $\tilde{S}$-matric is somehow constructed and we have it in our hands ready for use, the whole non-physical subspace becomes irrelevant; we have a theory which works in the physical subspace $\mathscr{H}_p$ with positive norm only. It was shown [53], for example, that it is possible to construct a complicated hermitian operator from the projection operators and from $H(t)$ and $U(t)$ and its inverse, such that if it were substituted for $W(t)$ in (5.3), through an always norm-conserving process, just the physical $\tilde{S}$-matrix (5.9) could be obtained. The presence of operators of the type $P_p U(t) P_p$ shows that here an (extremely complicated) non-local theory is obtained. Thus such modifications become equivalent to non-local field theories, in which, however, the metric is positive definite [19]. Cfr. [48].

Non-local field theories usually lead to non-causality. Indeed, Bogoljubov's original assumption was confirmed by SLAVNOV and SUKHANOV [55]. They have proved that even the « weak » causality condition

$$\tilde{S}(g_2 + g_1) \sim \tilde{S}(g_2)\ \ \tilde{S}(g_1)$$

cannot be ensured for those proposals, if the following requirements are made:

(i) relativistic invariance;

(ii) unitarity: $\tilde{S}\tilde{S}^* = P_v$;

(iii) $\tilde{S}$ should be constructed of S and the P-s.

In (5.10) $g(x)$ measures the strength of the interaction, g_1 and g_2 means that the interaction is switched on for the space-time regions Γ_1 and Γ_2: The asymptotic equality (weak causality) should be valid at least if $\Gamma_2 \gg \Gamma_1$, *i.e.* if the region in which g_2 is different from zero is located far later in time, than that where g_1 is different from zero.

The reason why it is impossible to obtain a causal physical scattering matrix is, essentially, that the physical part of a state vector is only a distinguished component of the whole vector, and although the change in time of the latter is a causal process, that of the distinguished component alone seems to be not causal.

6. – Conclusion.

In this paper we intended to give a short summary of the results obtained in quantum field theories with an indefinite metric. We wished to give here a complete list of references published until the first half of the year 1960. The author wants to apoligise if some paper has not been mentioned due to the fact that he did not have the opportunity to get a copy of it.

We have seen that, except quantum electrodynamics, where the additional Lorentz condition eliminates the difficulties, only theories with a pole type of ghost are physically acceptable. The other theories are not immediately probabilistically interpretable, and attempts to eliminate this obstacle have run into difficulties in connection with the causality principle. Therefore if this line proves a fruitful one at all, further results can be expected mainly from theories with pole ghosts. According to the considerations given especially in Sect. 4·1 it is not excluded, that in this way a finite and at the same time probabilistically interpretable field theory can be obtained. From this point of view Heisenberg's theory intends to follow just this line.

* * *

The author is greatly indebted to Professor B. TOUSCHEK for his interest, encouragement and helpful criticism and to the Scuola di Perfezionamento in Fisica Nucleare dell'Università di Roma for financial support.

REFERENCES

[1] P. A. M. Dirac: *Comm. Dublin Inst. Adv. Studies*, A, no. 1 (1943).
[2] W. Pauli: *Rev. Mod. Phys.*, **15**, 175 (1943).
[3] S. N. Gupta: *Proc. Phys. Soc.*, A **63**, 681 (1950).
[4] K. Bleuler: *Helv. Phys. Acta*, **23**, 567 (1950).
[5] S. N. Gupta: *Proc. Phys. Soc.*, A **64**, 695 (1951).
[6] S. N. Gupta: *Proc. Phys. Soc.*, A **65**, 161, 608 (1952).
[7] S. N. Gupta: *Proc. Phys. Soc.*, A **66**, 129 (1953).
[8] A. Pais and G. E. Uhlenbeck: *Phys. Rev.*, **79**, 145 (1950).
[9] P. T. Matthews: *Proc. Cambr. Phys. Soc.*, **45**, 441 (1949).
[10] R. B. N. Phillips: *Nuovo Cimento*, **1**, 823 (1955).
[11] T. W. B. Kibble and J. C. Polkinghorne: *Nuovo Cimento*, **10**, 74 (1958).
[12] K. L. Nagy and J. Rzewuski: *Bull. Acad. Polon. Sci.*, **7**, 93 (1959).
[13] T. D. Lee: *Phys. Rev.*, **95**, 1329 (1954).
[14] G. Källén and W. Pauli: *Mat. Fys. Medd.* **30**, no. 7 (1955).
[15] W. Heisenberg: *Nucl. Phys.*, **4**, 532 (1957).
[16] R. Ascoli and E. Minardi: *Nuovo Cimento*, **14**, 1254 (1959).
[17] W. Heisenberg: *Rev. Mod. Phys.*, **29**, 269 (1957); *Zeits. Naturfor.*, **14** a, 441 (1959) other reference here.
[18] N. N. Bogoljubov, B. V. Medvedev and M. K. Plivanov: Dubna preprint no. P 176 (1958); *CERN Conference Reports*, 129 (1958), *Sci. Reports of the Higher School, Phys. Mat. Series*, no. 1 (1958).
[19] K. L. Nagy: *Nuovo Cimento*, **10**, 1071 (1958).
[20] B. V. Medvedev and M. K. Polivanov: Dubna preprint no. P 184 (1958).
[21] M. A. Markov: *Nucl. Phys.*, **10**, 140 (1959).
[22] M. A. Markov and A. A. Komar: *Nucl. Phys.*, **12**, 190 (1959).
[23] M. Froissart: *Suppl. Nuovo Cimento*, **14**, 197 (1959).
[24] S. N. Gupta: *Canad. Journ. Phys.*, **35**, 961 (1957).
[25] L. K. Pandit: *Suppl. Nuovo Cimento*, **11**, 157 (1959).
[26] G. Konisi and T. Ogimoto: *Prog. Theor. Phys.*, **20**, 868 (1958); **21**, 727 (1959).
[27] S. T. Ma: *Phys. Rev.*, **75**, 535 (1949); F. J. Belinfante: *Phys. Rev.*, **76**, 226 (1949)
[28] S. Machida: *Progr. Theor. Phys.*, **14**, 407 (1955).
[29] J. S. Goldstein: *Nuovo Cimento*, **9**, 504 (1958).
[30] J. Lopuszanski: *Physica*, **25**, 745 (1959), other references here.
[31] E. V. Teodorovich: *D.A.N. SSSR*, **126**, 1236 (1959).
[32] K. L. Nagy: *Nuovo Cimento* **17**, 384 (1960).
[33] B. Birula: *Bull. Acad. Polon. Sci.*, **7**, 77 (1959).
[34] W. E. Thirring: *Ann. Phys.*, **9**, 91 (1958); V. Glaser: *Nuovo Cimento*, **9**, 990 (1958).
[35] G. Wentzel: *Helv. Phys. Acta*, **15**, 111 (1942).
[36] L. D. Landau and I. Pomeranchuk: *D.A.N. SSSR*, **102**, 489 (1955); I. Pomeranchuk: *Nuovo Cimento*, **3**, 1186 (1956); H. Umezawa, Y. Tomozawa, M. Konuma and S. Kamefuchi: *Nuovo Cimento*, **3**, 772, 1060 (1956); S. Kamefuchi: *Mat. Fys. Medd.*, **31**, no. 6 (1957).
[37] G. Källén: *CERN Symposium*, **2**, 187 (1956); *CERN Publications*, T/GK 3, (1955); 57-43 (1957); *Suppl. Nuovo Cimento*, **14**, 105 (1959).
[38] K. W. Ford: *Phys. Rev.*, **105**, 320 (1957).

[39] G. Källén: *Helv. Phys. Acta*, **25**, 417 (1952); H. Lehman: *Nuovo Cimento*, **11**, 342 (1954).
[40] K. L. Nagy: *Nuovo Cimento*, **15**, 993 (1960).
[41] P. J. Redmond: *Phys. Rev.*, **112**, 1404 (1958); *Phys. Rev. Lett.*, **1**, 398 (1958).
[42] P. J. Redmond and J. L. Uretsky: *Phys. Rev. Lett.*, **1**, 147 (1958).
[43] N. N. Bogoljubov, A. A. Logunov and D. V. Skirkov: *Žurn. Eksp. Teor. Fiz.* **37**, 805 (1959).
[44] B. V. Medvedev and M. K. Polivanov: *D.A.N. SSSR*, **127**, 537 (1959); T. Ogimoto and Y. Yamamoto: *Progr. Theor. Phys.*, **23**, 218 (1960).
[45] A. Uhlmann: *Nucl. Phys.*, **9**, 588 (1959).
[46] A. Uhlmann: *Nucl. Phys.*, **12**, 103 (1959).
[47] R. Ascoli and E. Minardi: *Nuovo Cimento*, **8**, 951 (1958).
[48] R. Ascoli and E. Minardi: *Nucl. Phys.*, **9**, 242 (1958).
[49] W. Heitler: *Quantum Theory of Radiation*, 3rd ed., App. 3 (Oxford, 1955).
[50] V. G. Vaks: *Žurn. Eksp. Teor. Fiz.*, **37**, 467 (1959).
[51] H. Mitter: *Zeits. Naturfor.*, **159**, 753 (1960).
[52] B. Ferretti: *Nuovo Cimento*, **12**, 393 (1959).
[53] K. L. Nagy: *Acta Phys. Hung.*, **11**, 193 (1960).
[54] L. A. Maksimov: *Žurn. Eksp. Teor. Fiz.*, **36**, 465 (1959).
[55] D. A. Slavnov and A. D. Sukhanov: *Žurn. Eksp. Teor. Fiz.*, **36**, 190 (1959)

SUPPLEMENTO AL VOLUME XVII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1960
3° Trimestre

Experimental and Theoretical Remarks on the Double β-decay.

G. F. DELL'ANTONIO (*)

Physics Department, Northwestern University - Evanston, Ill.

E. FIORINI (**)

Istituto di Fisica dell'Università - Milano
Istituto Nazionale di Fisica Nucleare - Sezione di Milano

(ricevuto il 3 Agosto 1960)

CONTENTS. **Introduction. – 1. Outline of the theory of β-decay. – 2. Theoretical aspects of the double β-decay. – 3. Experimental results on the double β-decay.** 1. Counter experiments. 2. Cloud chamber experiments. 3. Nuclear emulsions experiments. 4. Experiments with chemical methods. **4. Mass and nuclear structure of suitable isotopic triplets.** 1. Informations about the energy of the nuclear states. 2. Nuclear structure. 3. Possible double β-transitions. – **Conclusions.**

Introduction.

Renewed interest has been raised during the last few years in the process of the double β-decay, which is sometimes regarded as fairly important for the understanding of the structure of weak interactions. No sound conclusion can be drawn so far from the comparison of the theoretical predictions with the experimental results. Amusingly enough, most of the sources of uncertainty have nothing to do with what one would like to know, namely the form and the properties of the Hamiltonian responsible for weak interactions.

It is well known that the double β-decay consists of the spontaneous disintegration of a nucleus (A, Z) in the nucleus $(A, Z \pm 2)$. This transition

(*) On leave of absence from the Istituto Nazionale di Fisica Nucleare, Sezione di Milano; supported in part by a N.S.F. grant.

(**) On leave of absence at the Physics Department, Duke University, Durham, N. C.

is accompanied by the simultaneous emission of two electrons (or two positrons) with or without emission of neutrinos; which of the two alternatives takes place in nature is precisely the problem we are interested in. The transition $(A, Z) \to (A, Z-2)$ can occur also through a double electron capture or through a single electron capture and a single positron emission, again with or without emission of neutrinos.

A large number of experiments, stimulated by the well known neutrino-antineutrino puzzle (*i.e.*, whether or not they are different entities), were carried out whit different techniques. The decreasing values established as maximum activities for double transitions were often regarded as an experimental evidence in favor of the existence of physically distinguishable neutrino and antineutrino. The situation has become more complicated in recent years, expecially after the discovery of the non-conservation of parity in weak interactions, due to new suggestions which have been made (or revived) as to the nature of the neutrino and the structure of the Fermi Hamiltonian. As we shall see presently, experiments on the double β-decay are, by themselves, not sufficient to discriminate between the different possibilities.

The present authors believe that a general review of the problem of double β-decay with some theoretical and experimental remarks and suggestions on the more important points should not be completely useless for the future improvement of searchs on this phenomenon.

In the first part of this paper we have summarized those main features of the Fermi interaction that seem important for a clear understanding of the theoretical problems connected with the double β-decay. We have then discussed the phenomenon of the double β-decay both from the theoretical and from the experimental point of view, taking into particular account the possible improvements of the present results. Problems connected with the mass and the structure of the nuclei, and the overall situation, as it appears to the present authors, are briefly presented in the last part of this paper.

1. – Outline of the theory of β-decay.

Weak interactions can probably be fairly well described, at least in the short range of energy which is relevant for the decay of nuclei, by the phenomenological Fermi Hamiltonian density.

$$
(1.1)\quad \begin{cases} H_{\mathrm{int}}(x) = \sum_{i}^{\mathrm{II}} (\overline{\psi}_{\mathrm{n}}(x) O_i \psi_{\mathrm{p}}(x)) \left[f_{\mathrm{I},i}(\overline{\psi}_{\nu} O_i \psi_{\mathrm{e}}) - f'_{\mathrm{I},i}(\overline{\psi}_{\nu} \gamma_5 O_i \psi_{\mathrm{e}}) + f_{\mathrm{II},i}(\psi_{\nu} C O_i \psi_{\mathrm{e}}) + \right. \\ \qquad\qquad \left. + f'_{\mathrm{II},i}(\psi_{\nu} C \gamma_5 O_i \psi_{\mathrm{e}}) \right] + \mathrm{h.c.}\,, \\ O_1 = 1;\ O_2 = \gamma_5;\ O_3 = \gamma_\mu;\ O_4 = \gamma_\mu \gamma_5;\ O_5 = \tfrac{1}{2}(\gamma_\nu \gamma_\mu - \gamma_\mu \gamma_\nu)\,, \end{cases}
$$

$\psi^c(x) = C^{-1}\overline{\psi}(x)$ represents the solution to the Dirac equation charge-conjugate with respect to $\psi(x)$; we recall that the matrix C has the property $\gamma_\mu^T = -C\gamma_\mu C^{-1}$. The $\overline{\psi}_i(x)$ are operators which create a particle of type « i » or destroy the corresponding antiparticle. The f's are coefficients which characterize the strength of the interaction; their dimension is the square of a length, in units in which $\hbar = c = 1$. We use here the notation of [1].

There is no *a priori* reason why the neutrino and the antineutrino should be different entities. If they are not, namely if $\psi_\nu(x) = C^{-1}\overline{\psi}_\nu(x)$, the neutrino will be referred to as a Majorana particle. All coefficients in equation (1.1) may, in principle, be different from zero. We assume invariance of the theory under time-reflection so that all coefficients can be taken real.

The possibility of an interaction through an intermediate vector boson has been presented many times [2]. The conclusions we shall draw with regard to the double β-decay would not be changed in their essence if the formulation with an intermediate vector boson were chosen instead of the « direct » one; we shall therefore restrict ourselves, from now on, to the latter.

We shall assume that the mass, m_ν, of the neutrino is zero, and mention briefly the consequences of $m_\nu \neq 0$, when relevant. It is easily seen that, if primed and unprimed coefficents are both different from zero, non-conservation of parity follows. In those experiments in which no effort is made to distinguish between neutrinos and antineutrinos, the condition under which no parity-violation effect is observed is slightly weaker [3]. If coefficients with index I and II are both present, the number of leptons is not conserved.

Instead of dwelling on the different results one should expect from the many possible choices of coefficients, we shall concentrate on three particularly symmetric ones, which appear to be most relevant.

a) $f_{II} = f'_{II} = 0$, $\nu \neq \bar{\nu}$ (This is a short-hand notation for $\psi_\nu(x) \neq C^{-1}\overline{\psi}_\nu(x)$). The neutrino is a Dirac particle and leptons are conserved. This means that if to each lepton (ν, e^-, ...) is attributed a unit of an additive charge N_1, and to each antilepton minus one unit, the total charge N_1 is the same in the initial and final states of all physical processes.

b) $\nu = \bar{\nu}$. Lepton conservation is here certainly violated. It is worth emphasizing that condition *b*) is meant to mean: $P(I', F) = \exp[i\varphi]P(I', F')$ where $P(I, F)$ is the transition amplitude from a state I to a state F, φ an arbitrary phase factor, I' and F' states which differ from I, and F respectively only because, in their construction, neutrinos have been substituted for anti-neutrinos, and conversely.

c) The neutrino is a « Weyl » particle: *i.e.* the one-neutrino state is described by an irreducible representation of the Lorentz group which contains (only once) a Weyl representation. This would formally correspond to the

case $f_{\mathrm{I}} = f_{\mathrm{II}} = f$; $f'_{\mathrm{I}} = -f'_{\mathrm{II}} = f'$ [4-6]. This possibility arises when $m_\nu = 0$; in this case the solution to the Dirac equation has not only a representation through a spinor with four components, as appropriate to the Dirac and Majorana cases, but a representation whith two components or Weyl-representation as well. Two possible states are associated with a definite momentum; they have opposite helicity (*) and are interchanged by the operation of « charge » conjugation. They describe therefore a particle-antiparticle pair (which one is chosen to be the particle is of course a matter of convention). The theory is not invariant for a inversion of the space axes; in fact, this operation cannot even be defined (applied to the state of a neutrino it would give a neutrino state with « wrong » helicity, which does not exist in the framework of the theory). In this formulation of the Weyl theory conservation of leptons is ensured, and, in particular, the transition from a state which contains no leptons to a state which contains two electrons and no neutrinos is forbidden ($f' = 0$). If another solution to the Weyl equation is introduced, the helicity state previously called neutrino being now called antineutrino (and conversely), and the four resulting states are arranged in a suitable way, case *a*) is reproduced. Inversion of space axis now leads from one to the other type of neutrino. If moreover neutrino « one » is identified with antineutrino « two » (thus introducing a maximal non conversation of lepton number), the Majorana case is reproduced. This statement is sometimes referred to as (formal) equivalence of the two formulations.

We would like to present a survey of a very elegant treatment [1, 3, 7-8] of processes due to weak interactions which, although not sufficient to derive quantitative results, exhibits most of the qualitative features in a rather compact form. Consider the canonical transformation

$$\psi'(x) = a_1\psi(x) + b_1\gamma_5 C^{-1}\overline{\psi}(x) + a_2\gamma_5\psi(x) + b_2 C^{-1}\overline{\psi}(x)\,, \tag{1.2}$$

where

$$|a_1|^2 + |a_2|^2 + |b_1|^2 + |b_2|^2 = 1\,,$$

$$a_1 b_2 - b_2 a_1 = 0\,, \qquad \mathrm{Re}\,(a_1 a_2^*) + \mathrm{Re}\,(b_1 b_2^*) = 0\,.$$

If the mass of the particle represented by the field is zero, this transformation leaves the corresponding free field Hamiltonian invariant. We shall therefore use this transformation for the neutrino field.

An elementary computation shows that the canonical transformation is equivalent, for the interaction Hamiltonian (1.1), to the following linear

(*) The helicity is ± 1 according to whether the spin is parallel or antiparallel to the momentum. No other direction of spin is available for a particle with zero mass.

transformation of the coupling constants:

$$
(1.3)\quad \begin{cases} f_{\mathrm{I},i} \to a\ \cos\alpha f_{\mathrm{I},i} + b\ \cos\alpha f'_{\mathrm{II},i} + ia^*\sin\alpha f_{\mathrm{II},i} + ib^*\sin\alpha f'_{\mathrm{I},i}\ , \\ f_{\mathrm{II},i} \to a^*\cos\alpha f_{\mathrm{I},i} + b^*\cos\alpha f'_{\mathrm{I},i} + ib^*\sin\alpha f_{\mathrm{I},i} - ia^*\sin\alpha f_{\mathrm{II},i}\ , \\ f'_{\mathrm{I},i} \to a^*\cos\alpha f_{\mathrm{II},i} + b^*\cos\alpha f'_{\mathrm{I},i} + ia\ \sin\alpha f_{\mathrm{I},i} + ib\ \sin\alpha f'_{\mathrm{II},i}\ , \\ f'_{\mathrm{II},i} \to a^*\cos\alpha f_{\mathrm{II},i} + b^*\cos\alpha f'_{\mathrm{I},i} - ia^*\sin\alpha f'_{\mathrm{II},i} + ib^*\sin\alpha f_{\mathrm{I},i}\ , \\ a \equiv \sqrt{a_1^2 - a_2^2}\ , \qquad b \equiv \sqrt{b_1^2 - b_2^2}\ , \qquad \alpha = -\operatorname{arctg}\left(-\frac{ia_2}{a_1}\right) \end{cases}
$$

Introducing the combinations

$$
\begin{aligned} F_{\mathrm{I},i} &\equiv f_{\mathrm{I},i} - f'_{\mathrm{I},i}\ , & G_{\mathrm{I},i} &= f_{\mathrm{I},i} + f'_{\mathrm{I},i}\ , \\ F_{\mathrm{II},i} &= f_{\mathrm{II},i} - f'_{\mathrm{II},i}\ , & G_{\mathrm{II},i} &= f_{\mathrm{II},i} + f'_{\mathrm{II},i}\ , \end{aligned}
$$

the linear transformation (1.3) takes the form

$$
(1.3')\quad \begin{cases} F_{\mathrm{I},i} \to a\exp[-i\alpha]F_{\mathrm{I},i} + b\exp[-i\alpha]F_{\mathrm{II},i}, \\ G_{\mathrm{I},i} \to a\exp[i\alpha]G_{\mathrm{I},i} + b\exp[i\alpha]G_{\mathrm{II},i}\ , \\ F_{\mathrm{II},i} \to -b^*\exp[i\alpha]F_{\mathrm{I},i} + a^*\exp[-i\alpha]F_{\mathrm{II},i}, \\ G_{\mathrm{II},i} \to -b^*\exp[i\alpha]G_{\mathrm{I},i} + a^*\exp[i\alpha]G_{\mathrm{II},i}\ . \end{cases}
$$

It is rather easy to check the invariance of the quantities

$$
(1.4)\qquad F_{\mathrm{I}}F_{\mathrm{I}}^* + F_{\mathrm{II}}F_{\mathrm{II}}^*; \qquad G_{\mathrm{I}}G_{\mathrm{I}}^* + G_{\mathrm{II}}G_{\mathrm{II}}^*; \qquad F_{\mathrm{I}}G_{\mathrm{II}} - F_{\mathrm{II}}G_{\mathrm{I}}\ ,
$$

under (1.3). It is also easily seen that, in the group of linear transformations (not including therefore complex conjugation) only the subgroup (1.3) leaves invariant the quantities (1.4). The invariance of (1.4) characterizes therefore completely the trasformation (1.3).

We now recall that, if a canonical transformation leaves the initial and final states invariant, the transition amplitudes must also be invariant, apart from a common phase factor. In our case this implies that the transition between states which are left invariant by the canonical transformation (1.2) can only depend on those combinations of the coefficients (1.3′) which are invariant up to a phase factor. Two combinations which acquire different phase factors cannot appear simultaneously. The three invariants (1.4), written in terms of f_{I} ... etc. take the form

$$
\begin{aligned} \tfrac{1}{2}(1^{\mathrm{st}} + 2^{\mathrm{nd}}) \equiv K_{ij} = K_{ji}^* &= f_{\mathrm{I}i}^* f_{\mathrm{I}j} + f_{\mathrm{I}i}'^* f'_{\mathrm{I}j} + f_{\mathrm{II}i}^* f_{\mathrm{II}j} + f_{\mathrm{II}i}'^* f'_{\mathrm{II}j}\ , \\ \tfrac{1}{2}(1^{\mathrm{st}} - 2^{\mathrm{nd}}) \equiv L_{ij} = L_{ji}^* &= f_{\mathrm{I}i}^* f'_{\mathrm{I}j} + f_{\mathrm{I}i}'^* f_{\mathrm{I}j} - f_{\mathrm{II}i}^* f'_{\mathrm{II}j} - f_{\mathrm{II}i}'^* f_{\mathrm{II}j}\ . \end{aligned}
$$

Simmetric part of $3^{\mathrm{rd}} \equiv I_{ij} = I_{ji} = f_{\mathrm{I}i}f_{\mathrm{II}j} + f_{\mathrm{I}j}f_{\mathrm{II}i} + f'_{\mathrm{I}i}f'_{\mathrm{II}j} + f'_{\mathrm{II}i}f'_{\mathrm{I}j}$

Antisymm. part of $3^{\mathrm{rd}} \equiv J_{ij} = -J_{ji} = f_{\mathrm{I}i}f'_{\mathrm{II}j} - f_{\mathrm{I}j}f'_{\mathrm{II}i} + f'_{\mathrm{I}i}f_{\mathrm{II}j} - f'_{\mathrm{I}j}f_{\mathrm{II}i}$.

For completeness we list here below the invariants up to a phase factor and their transformation properties

$$M_{ij} \equiv F^*_{\mathrm{I}i} G_{\mathrm{I}j} + F^*_{\mathrm{II}i} G_{\mathrm{II}j} \qquad M_{ij} \to M_{ij} \exp[-2i\alpha],$$

$$N_{ij} \equiv F_{\mathrm{I}i} F_{\mathrm{II}j} - F_{\mathrm{II}i} F_{\mathrm{I}j} = -N_{ji} \qquad N_{ij} \to N_{ij} \exp[-2i\alpha],$$

$$N^R_{ij} \equiv G_{\mathrm{I}i} G_{\mathrm{II}j} - G_{\mathrm{II}i} G_{\mathrm{I}j} = -N^R_{ji}, \qquad N^R_{ij} \to N_{ij} \exp[2i\alpha].$$

We notice that, if the neutrino is a Dirac particle and if lepton number is conserved, one has $f_{\mathrm{I},i} = f'_{\mathrm{II},i} = 0$. In this case, therefore, $I = J = N = N^R = 0$. The Majorana case is obtained with the choice

$$f_{\mathrm{I}i} = -f_{\mathrm{II}i} = \frac{1}{\sqrt{2}} f_i; \qquad f'_{\mathrm{I}i} = f''_{\mathrm{II}i} = \frac{1}{\sqrt{2}} f'_i,$$

therefore

$$I_{ij} = f'_i f'_j - f_i f_j; \quad J_{ij} = f_i f'_j - f'_i f_j; \quad K_{ij} = f^*_i f_j + f'^*_i f'_j; \quad L_{ij} = f^*_i f'_j + f'^*_i f_j.$$

As already noted the two components neutrino theory appears formally equivalent to the Majorana case if maximal non-conservation of lepton number is postulated; if conservation of lepton number is required one has $f_i = f'_i$ or $f_i = -f'_i$, the same sign holding for all « i » if interacting neutrinos are to be regarded as the same particle (*). Therefore in this case too

$$I_{ij} = J_{ij} = N_{ij} = N^R_{ij} = 0.$$

Let us finally mention the situation in a theory [13-14] in which the neutrino is treated as a Weyl particle but the Hamiltonian is slightly different from (1.1). We whall consider only interaction between nucleons, mesons and leptons, and forget about strange particles. The Hamiltonian density is written as [13, 15-17]

$$(1.5) \quad \begin{cases} H_{\mathrm{int}}(x) = \left(J^-_{\mathcal{N}}(x) + J_{\mathrm{l}}(x)\right)\left(J_{\mathcal{N}}(x) + J_{\mathrm{l}}(x)\right), \\ J_{\mathrm{l}} = \left(\overline{\psi}_{\nu}(x)\gamma_\mu(1-\gamma_5)\psi_{\mathrm{l}}(x)\right); \quad J_{\mathcal{N}}(x) = \left(\overline{\psi}_{\mathrm{n}}(x)\gamma_\mu(1-(1+\alpha)\gamma_5)\psi_{\mathrm{p}}(x)\right) + \\ \qquad + \left(\boldsymbol{\Phi}(x) \wedge \partial_\mu \boldsymbol{\Phi}(x)\right)_+ + \frac{m}{g}\partial_\nu \boldsymbol{\Phi}_+(x), \end{cases}$$

(*) The choice $f_i = f'_i$ for some values of the index and $f_i = -f'_i$ for others has been considered by various authors [9-12]; this corresponds to the case in which the neutrinos coming from the scalar, pseudoscalar, vector... parts of the Hamiltonian have different properties.

where g is the strong coupling constant in the *ps-ps* theory; $\mathbf{\Phi}(x)$ is the meson field operator (it is bold face to signify that it is a vector in isotopic spin space). For a generic vector $\boldsymbol{A}$ the notation A_+ stands for $\frac{1}{2}(A_1 + iA_2)$, being $A \equiv (A_1, A_2, A_3)$; $\mathbf{\Phi}_+$ is, in particular, the operator which creates a π^+ or destroys a π^-.

The factor α is regarded as a renormalization factor and has to be introduced in the fenomenological Hamiltonian density (1.5) because the axial part of J_λ is exactly conserved only in the limit of zero meson mass. Since the divergence of J_λ (axial) is proportional to m_π one might expect α to be proportional to $m_\pi/m_n \simeq .15$. Notice that there is no such factor in the leptonic part since we have assumed $m_\nu = 0$; we remind here that, within the framework of pertubation theory, if the bare mass of a particle is zero, its true mass is also zero.

One can see from (1.5) that the emission of an electron must be accompanied by the emission of an antineutrino or the absorption of a neutrino; the emission of two electrons and no neutrino is forbidden. The same conclusion can be reached, of course, using the formulation in terms of invariants. If $m_\nu \neq 0$, one should expect a factor $1 + \alpha'$, with α' possibly much smaller than α, due to the different strength of the renormalizing interactions. In this case $(1 - \gamma_5)\psi_\nu$ is no longer an eigenstate of the free neutrino Hamiltonian and there is a finite probability to find the neutrino after a finite time in a state of helicity opposite to the original one (the four components theory has to be used here). To this probability contribute both the free particle Hamiltonian, through the mass term, and the interaction Hamiltonian, through the term containing α'. Such probability is therefore expected to contain three terms, proportional to m_ν^2, α'^2, $|m_\nu \alpha'|$, respectively, the probability being proportional to the square modulus of the transition amplitude (*). We want however to stress again that, if $m_\nu = 0$, the Feynman-Gell Mann Hamiltonian forbids the emission of two electrons without neutrinos.

2. – Theoretical aspects of the double β-decay.

We are going to consider in this part a particular decay from a nucleus to one of its isobars through weak interactions. We assume, as usual, that the transition amplitude is given in terms of the Hamiltonians discussed in the previous part, although the initial and final states are here nuclear states. This amounts to assuming that the shell model is suitable for our purposes,

(*) For an explicit expression see E. GREULING and R. C. WHITTEN: *Lepton conservation and double β-decay*, preprint; we are very grateful to Dr. GREULING for letting us know his results prior to pubblication.

namely, that in nuclear states the single nucleons do not lose their individuality. We shall consider only phenomenological Hamiltonian density of the form (1.1); if one wants to use (1.5), corrections should be introduced arising from those terms which contain only the meson field. The qualitative features of the results described in this part are not sensitive to these corrections.

The process we are going to discuss can be characterized as follows: Let A, B, C, represent three nuclei with the same atomic number A_0 and charge $Z-1$, Z, $Z+1$, respectively. We shall indicate with μ_A, μ_B, μ_C their masses and with M_A, M_B, M_C, the masses of the corresponding atoms in their ground states. One has obviously $M = \mu + Zm_e - \eta$; η being the total binding energy of the system, and m_e the electron mass. If $M_A > M_B$, the decay $A \to B + e^- + \nu$ (*) is energetically allowed and so is $B \to C + e^- + \nu$ if $M_B > M_C$.

The decay in which we are interested, and which will be referred to as double β-decay is the process $A \to C +$ leptons (this implies $M_A > M_C$), when $M_A < M_B$, so that the first step in the sequence of two is a virtual one. Even if $M_A > M_B$ a virtual transition through an excited state of the atom B might dominate over the real one, due to selection rules. One relevant example is *e.g.* the transition ${}^{48}_{20}\mathrm{Ca} \to {}^{48}_{22}\mathrm{Ti}$ through ${}^{48}_{21}\mathrm{Sc}$ [18-19]; we shall come back briefly on this situation later on.

We also note that most of what will be said on the double β-decay is also true for the case when two positrons, rather than two electrons, are emitted (**). The relation between masses is; in this case, $M_A < M_B + 2m_e$; $M_A > M_C + 4m_e$. The same considerations apply to the process $(e^- + e^- + A)_{\text{bound}} \to C +$ leptons and/or γ-rays, when the transition $e^- + A \to B + \nu$ is forbidden (double electron capture), and to the process $e^- + A \to C + e^-$ plus possibly other leptons, when the transition $e^- + A \to B + \nu$ is forbidden (virtual electron capture followed by positive β-decay). The corresponding relations between masses are obvious, and will not be written explicitely; we note that the binding energy of the electrons plays here a slightly more important role, the electron which is absorbed being usually in a K shell. We also notice that if a process is energetically forbidden but for the presence of the binding energy in the final state, the corresponding transition amplitude is reduced by one (or more) factors α (fine structure constant), due to the « readjustment » of the electron cloud.

Let us remark that, if the inequalities we have written above for the atomic masses are satisfied, the transition $A \to C +$ leptons always takes place although

(*) The symbol ν stands here either for a neutrino or for an antineutrino.

(**) For the sake of simplicity, we always consider a chain $A \to B \to C$. In the case of the emission of two positrons, the charge assignement is $Z+1, Z, Z-1$ for A, B, C, respectively. Similar considerations hold for electron capture.

an exceedingly small rate can be anticipated, indipendently from any other factor, from the dependence of the transition probability on the fourth power of the very small Fermi coupling. The problem in which we are interested is to see whether a process of this kind can take place without emission of neutrinos or not; in the first the case, lepton conservation cannot be a law of nature, since two electrons are emitted. Needless to say, if no decay is observed in which no neutrino is emitted, we cannot draw any conclusion regarding lepton conservation, since the process might be forbidden for other reason. If a neutrinoless process takes place, we have a transition from an initial state which does not contain neutrinos or antineutrinos (the atom A) to a final state which has the same property (the atom C and two electrons). These states are both left invariant by the canonical transformation considered in Sect. **1** and therefore the matrix elements and the transition probability can only depend on the invariants I_{ij}, K_{ij}, N_{ij}, N^R_{ij}. It is easily seen that the complex conjugate of the coupling constants does not enter into the description of the process. We conclude that such a transition cannot take place if the invariants are zero; particularly in the two components theory, or in a theory with maximal non-conservation of parity (*i.e.* with conservation of helicity). For other choices of the parameters in the interaction Hamiltonian the probability ranges from zero to a maximum given by the Majorana theory (one can easily check that the invariants reach their maximum value corresponding to the choice of $f_{\mathrm{I}}, f_{\mathrm{II}}$ which characterizes this theory).

It may be useful to sketch a physical picture of these considerations; regard again the process $\mathrm{A} \to \mathrm{C} + \text{leptons}$ as a chain

$$\mathrm{A} \to \mathrm{B} + \nu_1 + \mathrm{e}^-; \qquad \mathrm{B} \to \mathrm{C} + \nu_2 + \mathrm{e}^-,$$

or

$$\mathrm{A} \to \mathrm{B} + \nu_1 + \mathrm{e}^-; \qquad \nu_1 + \mathrm{B} \to \mathrm{C} + \mathrm{e}^-,$$

where ν_1 and ν_2 are certain superpositions of states of neutrino and antineutrino, characteristic of the interaction chosen. If $P_{12} \equiv \langle \nu_1 | \bar{\nu}_2 \rangle \neq 0$ the transition $\mathrm{A} \to \mathrm{C} + \mathrm{e}^- + \mathrm{e}^-$ is possible, and its probability is proportional to $|P_{12}|^2$. Clearly $|P_{12}|^2 \leqslant 1$, and in particular $|P_{12}|^2 = 1$ when $|\bar{\nu}_2\rangle = \exp[i\alpha]|\nu_1\rangle$, namely when no distinction is made between neutrinos and antineutrinos. If, on the opposite, $|\nu_1\rangle$ represents a state of a zero-mass particle in which momentum and spin are parallel, and $|\bar{\nu}_2\rangle$ an analogous state in which they are antiparallel, conservation of momentum and angular momentum implies $\langle \bar{\nu}_2 | \nu_1 \rangle = 0$ and the only possible process is $\mathrm{A} \to \mathrm{C} + \mathrm{e}^- + \mathrm{e}^- + \nu + \bar{\nu}$.

This is the case in the two-component theory of Feynman-Gell Mann [13]; it is obviously the case also when the neutrino is a Dirac particle and, in the theory, conservation of leptons is postulated. Although not sufficient to discriminate between all possible choices of the coupling constants, experiments on the β-decay with evidence in favor of a neutrinoless ββ decay can rule out conser-

vation of number of leptons. This can be achieved also with the experiments [20] on the strictly correlated [7, 18, 21] problem of the absorption of neutrinos from a pile. The neutrinos come here from the decay of a sample of a known element. From a theoretical point of view this process differs from the ββ decay only because the neutrino which is absorbed is not a virtual one. We may note, incidentally, that the absorption of a real neutrino is possible also in the double β-decay, but the probability for a nucleus B to capture the neutrino emitted in the transition $A \to B + e^- + \nu_1$ is extremely small and the contribution of this process negligible. This is the reason why a double β-decay whithout neutrinos can be expected only if the first transition is virtual.

Let us notice that, when $|P_{12}|$ is not too small compared to unity, the transition without neutrinos is more likely to happen than the transition with two neutrinos. We shall give the explicit forms of the transition probabilities for both cases and see that the neutrinoless process is favored by a factor of about $10^5 \div 10^6$. Roughly speaking, the reason for this is the following: in the no-neutrino case a sum has to be taken of the contributions coming from virtual neutrinos with arbitrarly high energy. We shall presently see that it is rather reasonable to assume that the most important contribution comes from energies around 30 MeV. Very elementary considerations on the volume available in phase space lead then to a ratio for the transition probability of about 10^{-5}. We see therefore that if $|P_{12}|^2$ is not less than, say, 10^{-3}, almost all double β-decays are expected to go through the no-neutrino channel.

There is an experiment, very appealing from the theoretical point of view, which can easily reveal the presence of a decay without neutrinos. In such type of decay the difference in energy $M_A - M_B$ goes almost completely to the electrons, the energy taken away by the recoil of the nucleus C being negligible even in the most favorable distribution. If the sum $\varepsilon + 2m_e$ of the energies of the electrons can be measured and is plotted against the number of events, a single line (*) has to be observed at $\varepsilon_0 = M_A - M_C$ (or a sequence of lines if many excited states of the nucleus C can be reached). This experiment, the result of which could be compared with the only clean-cut theoretical result available, is unfortunately very hard to perform. A study of the most relevant difficulties will be more properly given in Sect. 3, devoted to a discussion of the relative merits of the different experimental techniques. We just mention here that the main difficulty lies in the extremely rare occurrence of the double decays; even if all the events had to take place within a small range of energies their frequency would be so low that it could hardly, if at all, be distinguishable from the background events in the same energy range.

(*) Or, more precisely, a sharp peak; the broadening is due to the recoil of the nucleus as well as to the width of the nuclear ground state.

In a process in which two neutrinos are emitted, these particles carry away an energy comparable with that of the electrons. The expression available for the distribution of the sum of the energies of the electrons are only roughly approximate, the approximation being presumably better [18] if

$$M_B - M_C \gg 1 - \frac{M_A - M_C}{2} \qquad \text{(units in MeV)} . \tag{2.1}$$

In this case the distribution of the total energy of the electrons is approximately given by the formula

$$P(\varepsilon)\,\mathrm{d}\varepsilon \sim (M_A - M_C - \varepsilon)^5 \varepsilon [\varepsilon^4 + 5\varepsilon^3 + 10\varepsilon^2 + 7.5\varepsilon + 2]\,\mathrm{d}\varepsilon , \tag{2.2}$$

where ε is the total energy of the two electrons minus the sum of their masses (Fig. 1). If the inequality (2.1) is not satisfied, the distribution might be, in particular, much shifted towards $M_A - M_C$, which is the maximum energy available to the electrons.

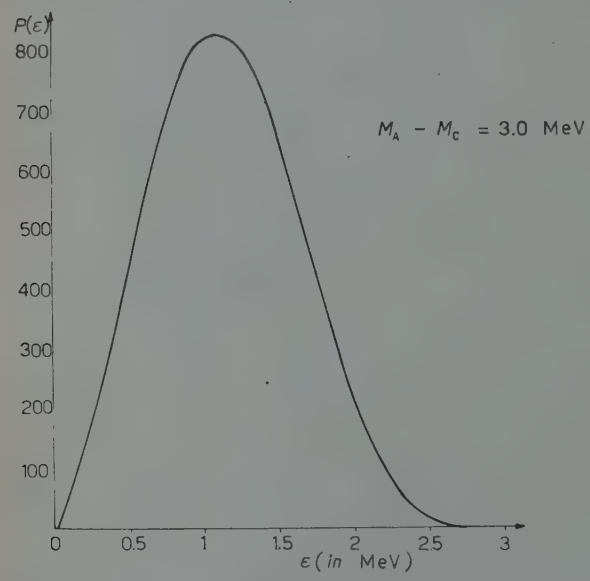


Fig. 1. – Theoretical distribution of the sum of the two electrons energies for double β-decay with emission of neutrinos.

The theoretical predictions for many other quantities, in particular the transition probability and therefore the lifetime, depend radically on the choice of the relative magnitude of the coupling constants. They could, therefore, be of use in trying to eliminate, by comparison with existing experimental data, choices otherwise acceptable. The precision and even sometimes the reliability of such theoretical data are however rather questionable, due to our poor knowledge of most of the nuclear states which come into play and to the fairly crude methods of approximation available to get reasonable expressions out of matrix elements of a rather simple Hamiltonian between complicated states. We do not attempt here to give a complete discussion of all the problems involved, neither to exploit in detail the reliability of each approximation. Fairly complete discussions are given, *e.g.*, in [18, 22-24]. We would rather give a general idea of the problems involved and quote some

results among those which seem to us most useful for an experimentalist interested in double β-decay.

The kind of difficulties which one has to deal with is most easily seen in the case of simple β-decay. The problem is here to compute the matrix element of the interaction Hamiltonian

$$H_{\text{int}} = \int \mathrm{d}^3x \, \mathcal{H}_{\text{int}}(x) \,,$$

between the initial state $|i\rangle$ of a nucleus and a final state of a nucleus plus a neutrino plus an electron, $|f\rangle$. The transition rate is then given in first order perturbation theory by

$$T \, \mathrm{d}E = 2\pi |\langle f | H_{\text{int}} | i \rangle|^2 \varrho(E) \, \mathrm{d}E \,,$$

$\varrho(E)$ being the density of the final states in the range $E \div E + \mathrm{d}E$. Expanding the neutrino operator in plane waves and the electron in wave functions [24] appropriate to the Coulomb field seen by the emitted electron, we have

$$(2.3) \quad \begin{cases} \psi_\nu(x) = \int \mathrm{d}^3\boldsymbol{q} \, \exp\left[i\boldsymbol{q}\cdot\boldsymbol{x} - i|\boldsymbol{q}|t\right] \sum_{1,2,j} a^+_{j,\nu}(\boldsymbol{q}) u^{j,+}_\nu(\boldsymbol{q}) + \\ \qquad + \int \mathrm{d}^3\boldsymbol{q} \, \exp\left[-i\boldsymbol{q}\cdot\boldsymbol{x} - i|\boldsymbol{q}|t\right] \sum_{1,2,j} a^-_{j,\nu} u^{j,-}_\nu(\boldsymbol{q}) \,, \\ \psi_e(x) = \int \mathrm{d}^3\boldsymbol{K} \Phi(Z, \varepsilon, \boldsymbol{K}\cdot\boldsymbol{x}) \sum_{1,2,j} a^+_{j,e}(\boldsymbol{K}, Z) u^{j,+}_e(\boldsymbol{K}) + \\ \qquad + \int \mathrm{d}^3\boldsymbol{q} \, \Phi(Z, \varepsilon, -\boldsymbol{K}\cdot\boldsymbol{x}) \sum_{1,2,j} \overset{*}{a}{}^-_{j,e}(\boldsymbol{K}, Z) u^{j,-}_e(\boldsymbol{K}) \,, \end{cases}$$

where (*) u^{j+}_α, $j = 1, 2$ are two indipendent, normalized positive-frequency solutions to the Dirac equation with mass appropiate to the particle α and u^{j-}_α, $j = 1, 2$ are the corresponding negative-frequency solutions. One gets

$$(2.4) \quad M = \langle f | H_{\text{int}} | i \rangle = \sum_i^5 \int \mathrm{d}^3x \langle f' | \psi_n(x) O_i \psi_p(x) \exp\left[-i\boldsymbol{q}\cdot\boldsymbol{x}\right] \Phi(Z, \varepsilon, \boldsymbol{K}\cdot\boldsymbol{x}) | i \rangle \cdot$$
$$\cdot \left\{ (\bar{u}_\nu(\boldsymbol{q})(f_{\mathrm{I},i} O_i - f'_{\mathrm{I},i} \gamma_5 O_i) u_e(\boldsymbol{K})) + (u_\nu(\boldsymbol{q})(f_{\mathrm{II},i} O_i + f_{\mathrm{II},i} \gamma_5 O_i) u_e(\boldsymbol{K})) \right\} .$$

The transitions rate has now the form

$$(2.5) \quad |M|^2 = \sum_i^5 \sum_j^5 \int\int \mathrm{d}^3x \, \mathrm{d}^3y \langle f' | \bar{\psi}_n(x) O_i \psi_p(x) \exp\left[-i\boldsymbol{q}\cdot\boldsymbol{x}\right] \Phi(Z, \varepsilon, \boldsymbol{K}\cdot\boldsymbol{x}) | i \rangle \cdot$$
$$\cdot \langle f' | \bar{\psi}_n(y) O_i \psi_p(y) \exp\left[-i\boldsymbol{q}\cdot\boldsymbol{y}\right] \Phi(Z, \varepsilon, \boldsymbol{K}\cdot\boldsymbol{y}) | i \rangle^* \cdot G_{n,m} \,,$$

(*) See, *e.g.* L. I. SCHIFF: *Quantum Mechanics* (New York, 1955), pp. 327.

where $G_{n,m}$ is a factor which contains only the leptonic part of the matrix element and can be treated exactly. The exponential in (2.4) can be expanded in Taylor series in $\boldsymbol{x}$; the transition rate is now the sum of terms which will be called of order « n » if they come from a term in the expansion that contains $(\boldsymbol{x})^n$. In general a term of order « n » is smaller, as compared with that of order zero, by a factor $K_{\max}^n \langle R \rangle^n$ where $K_{\max}$ is the maximum momentum available to the electron and $\langle R \rangle$ is the average nuclear radius. Due to the smallness of $K_{\max}$, this factor is small in all real β-decays. Notice, however, that in virtual β-decays, and more precisely in those decays in which the neutrino is a virtual particle, as is the case in both steps of the double β-decay without neutrinos, there is no upper bound to the energy of the neutrino, $|\boldsymbol{q}|$. The matrix element for a double β-decay contains, as is easily seen from (2.4) (*), a factor $\langle f | \exp[i\boldsymbol{q}(\boldsymbol{x}_m - \boldsymbol{x}_n)] | i \rangle$. Contributions from very high values of $|\boldsymbol{q}|$ are averaged to zero by the high number of oscillations of the factor within a region in which the nuclear matrix element is approximately constant. The main contribution in the integration over $\boldsymbol{q}$ comes therefore from the intermediate region, *i.e.* the region in which $|\boldsymbol{q}| \langle R_{mn} \rangle \sim 1$. $\langle R_{nm} \rangle$ is the average distance of two neutrons in the parent nucleus. This is the reason why, as already mentioned before, we expect that the most effective virtual neutrinos have an energy around 30 MeV. Going back to the matrix element for the single β-decay, we notice that, since the masses of the nucleons are much greater than the energy released, the nuclear matrix elements can be treated non-relativistically. In this approximation two of the components of the spin function $u_n^{\mu,\pm}(\boldsymbol{K})$ (2.3) for the nucleon are smaller than the remaining two by a factor of v/c; v being the velocity of the nucleon with momentum $\boldsymbol{K}$. The matrix elements in which two « large » components enter, are of the order of magnitude of the unity and are called ordinary. Those which contain one large and one small component are of order v/c and give rise to the so called relativistic transition of order zero [24].

With the above notations, allowed transitions are ordinary transitions of zero order; first forbidden are ordinary transitions of first order and relativistic transitions of zero order... . The nuclear matrix elements for ordinary transitions are obtained from those previously given carrying out in the integrands the following substitutions:

$$O_1 \equiv O_p \to 0; \quad O_2 \equiv O_s \to 1; \quad O_3 \equiv O_v \to 1; \quad O_4 \equiv O_T \to \boldsymbol{\sigma}; \quad O_5 \equiv O_A \to \boldsymbol{\sigma},$$

(*) Each matrix element contains $\int \exp[i\boldsymbol{q}\cdot\boldsymbol{x}]\cdot \mathrm{d}\boldsymbol{q}$ (we have already assumed that $\exp[i\boldsymbol{K}\cdot\boldsymbol{x}] \sim 1$). One gets therefore

$$\langle \int\int\int\int \exp[i\cdot\boldsymbol{q}_1\cdot\boldsymbol{x}_m] \exp[-i\boldsymbol{q}_2\cdot\boldsymbol{x}_n]\, \delta(\boldsymbol{q}_1 - \boldsymbol{q}_2) \rangle = \langle \int \exp\left[iq(\boldsymbol{x}_n - \boldsymbol{x}_m)\right] \mathrm{d}q\, \mathrm{d}x_n\, \mathrm{d}x_m \rangle,$$

the δ function coming, roughly speaking, from the fact that the same neutrino is emitted and absorbed.

($\boldsymbol{\sigma}$ is a vector which has, as components, the three Pauli matrices) and treating the initial and final states as assemblies of two-component spinors. The matrix elements S and V are called Fermi matrix elements, the A and T Gamow-Teller ones. The formers give transitions characterized by $\Delta I = 0$ ($I =$ nuclear spin), the latters by $\Delta I = \pm 1{,}0 \quad 0 \to 0$ being forbidden. Changes in spin larger than unity are possible only through first, or higher, forbidden transitions and the corresponding transition rates are consequentely smaller.

We now sketch briefly the derivation of the approximate forms of the transition amplitudes both with and without emission of neutrinos in double β-decay. Let us start from the latter. The matrix element we are considering is, neglecting electron clouds,

$$\langle \beta^-_{p_1,\mu_1} \beta_{p_2,\mu_2^*}(A, Z+2) \,|\, (A, Z)\rangle \,, \tag{2.6}$$

It gives the transition amplitude for the emission of two negative electrons with momenta p_1, p_2, and spin states characterized by μ_1, μ_2. The second order time-dipendent perturbation theory gives us

$$\langle \beta^-_{p_1\mu_1} \beta^-_{p_2\mu_2}(A, Z+2) \,|\, (A, Z)\rangle = \sum_{\text{int. states}} \langle \beta^-_{p_1\mu_1} \beta^-_{p_2\mu_2}(A, Z+2) \,|\, H_{\text{int}} \,.\, \beta^-_{p_2\mu_2} \nu_1(A, Z+1)\rangle \cdot \tag{2.7}$$

$$\cdot (E' - E_{\text{in}})^{-1} \cdot \langle \beta^-_{p_2\mu_2} \nu_1(A, Z+1) \,|\, H_{\text{int}} \,|\, (A, Z)\rangle \,,$$

where ν_1 is a certain superposition of neutrino and antineutrino states. E' and E_{in} represent the energies of the intermediate and final states respectively.

Expanding the intermediate neutrino and the intermediate and final electrons according to (2.3), symmetrizing with respect to electrons 1 and 2 and taking into account that, by considerations made in part I, the matrix element can only depend on the coefficients I_{ij}, J_{ij} we get [18], averaging over the intermediate neutrino states

$$\langle \beta^-_{p_1\mu_1} \beta^-_{p_2\mu_2}(A, Z+2) \,|\, (A, Z)\rangle = (1 - P_{12}) \sum_{i,j} \int \mathrm{d}^3\boldsymbol{K} \cdot \sum_{\text{int}} \cdot \tag{2.8}$$

$$\cdot \langle \psi_f \,|\, \sum_{n=1}^{A} \tau_n^+ O_n^i \Phi^*(Z, \varepsilon_2, \boldsymbol{p}_2 \cdot \boldsymbol{x}_n) \exp[i\boldsymbol{K} \cdot \boldsymbol{x}_n] \,|\, \psi_{\text{int}}\rangle \cdot (W_{\text{int}} - W_i + \varepsilon_1 + 1 - |K|)^{-1} \cdot$$

$$\cdot \langle \psi_{\text{int}} \,|\, \sum_{m=1}^{A} \tau_m^+ O_m^j \Phi(Z, \varepsilon_1, \boldsymbol{p}_1 \cdot \boldsymbol{x}_m) \exp[-i\boldsymbol{K} \cdot \boldsymbol{x}_m] \,|\, \psi_i\rangle \cdot$$

$$\cdot \left(u_e^{\mu_1 +}(\boldsymbol{p}_1) O^i \frac{1}{2} \left(1 + \frac{\boldsymbol{\alpha} \cdot \boldsymbol{K}}{|\boldsymbol{K}|} \right) C[I_{ij} + \gamma_5 J_{ij}] \tilde{O}^j u_e^{*\mu_2 +}(\boldsymbol{p}_2) \right).$$

In (2.8) τ_n^+ is an operator in the isotopic spin space which acts on the n-th nucleon, giving zero if it is in a « proton » state, and transforming it into a proton if it is in a neutron state. $|\psi\rangle$ represents the state of a nucleus and W its energy. $\frac{1}{2}(1 + \boldsymbol{\alpha} \cdot \boldsymbol{K}/|\boldsymbol{K}|)$ is the projection operator on the positive

energy oneparticle states of the virtual neutrino, or antineutrino. In a two-components neutrino theory $\boldsymbol{\alpha}$ should be replaced by $\boldsymbol{\sigma}$.

We have already seen that the major contribution comes from those virtual neutrinos for which $|K| = |K|_0 \simeq 30$ MeV. It easy to convince oneself that the corresponding recoil energy of the intermediate nucleus is smaller than $|K|_0$ by a factor of at least four; it will be therefore neglected in the denominator of the sum over the intermediate states obtained by closure. This approximation does not change the order of magnitude of the matrix element, and this is all what we are interested in. The term $\varepsilon_1 + 1$ is again small compared to $|K|_0$ and it will be neglected too. The integration over $\boldsymbol{K}$ is then simplified and we get

$$(2.9)\quad \langle \beta^-_{\boldsymbol{p}_1\mu_1}\beta^-_{\boldsymbol{p}_2\mu_2}(A, Z+2)\,|\,(A, Z)\rangle = \sum_{i,j}\frac{\pi^2}{(2\pi)^3}\,(u_e^{\mu_1+}(\boldsymbol{p}_1)O^iC[I_{ij}+\gamma_5 J_{ij}]\tilde{O}^j u_e^{*\mu_2,+}(\boldsymbol{p}_2))\cdot$$

$$\cdot\langle\psi_f|\sum_{n,m}\tau_n^+\tau_m^+ O_n^i O_m^j \Phi^*(Z, \varepsilon_2, \boldsymbol{p}_i\cdot\boldsymbol{x}_n)\,\Phi^*(Z, \varepsilon_2, \boldsymbol{p}_2\cdot\boldsymbol{x}_m)\cdot$$

$$\cdot\left[\frac{\varepsilon_2-\varepsilon_1}{|x_{nm}|} - 2i\,\frac{\boldsymbol{\alpha}\cdot\boldsymbol{x}_{nm}}{|x_{nm}|^3}\right]|\psi_i\rangle - (1\leftrightarrow 2) \qquad \boldsymbol{x}_{nm} = \boldsymbol{x}_n - \boldsymbol{x}_m\ .$$

We now take $\Phi(Z, \varepsilon, \boldsymbol{p}\cdot\boldsymbol{x}) \simeq \Phi(Z, \varepsilon, 0)(1 + i\boldsymbol{p}\cdot\boldsymbol{x})$ since the wave-length of the electron is large compared with nuclear dimension (*i.e.* $|\boldsymbol{p}_i|\langle|\boldsymbol{Z}|\rangle \ll 1$). For the sake of simplicity it is convenient to restrict ourselves to the very common case in which the nuclear states of the initial and final nuclei have both zero spin and the same parity. In this case only terms invariant for rotations and inversions of the nuclear coordinates survive, and one has

$$(2.9')\quad \langle \beta^-_{\boldsymbol{p}_1\mu_1}\beta^-_{\boldsymbol{p}_2\mu_2}(A, Z+2)\,|\,(A, Z)\rangle = \sum_{i,j}(u_e^{\mu_1+}(\boldsymbol{p}_1)O^iC[I_{ij}+J_{ij}\gamma_5]\tilde{O}^j u_e^{*\mu_2+}(\boldsymbol{p}_2))\cdot$$

$$\cdot\langle\psi_f|\left(\varepsilon_2-\varepsilon_1+\frac{\boldsymbol{\alpha}}{3}(\boldsymbol{p}_2-\boldsymbol{p}_1)\right)\sum_{m,n}\frac{\tau_n^+\tau_m^+ O_n^i O_m^j}{r_{mn}}\,|\psi_i\rangle\,\Phi^*(Z, \varepsilon_2, 0)\,\Phi(Z, \varepsilon_1, 0)\ .$$

Taking the square modulus of this expression, summing over all possible spin orientations of the final electrons and multiplying by the density of the final state one gets the probability distribution for the emission of two electrons with momenta $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ and no neutrino

$$P(\boldsymbol{p}_1, \boldsymbol{p}_2)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varepsilon_1\,\mathrm{d}\varepsilon_2 =$$

$$= \frac{(2\pi)^{-6}}{2}|p_1|\,|p_2|\,(\varepsilon_1+1)(\varepsilon_2+1)\,\delta(\varepsilon^0-\varepsilon_1-\varepsilon_2)\,\mathscr{F}(Z_1\varepsilon_1)\,\mathscr{F}(Z_2\varepsilon_2)\cdot$$

$$\cdot\left[A(\varepsilon_1\varepsilon_2) + B(\varepsilon_1\varepsilon_2)\cos\theta\right]\cdot\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varepsilon_1\,\mathrm{d}\varepsilon_2\ ,$$

where θ is the angle between $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, and ε^0 is the maximum energy available to the electrons. The δ-function is an expression of the fact that

if no neutrino is emitted the sum of the energies of the electrons must be the same in all processes. $\mathcal{F}(Z, \varepsilon)$, called Fermi factor, is equal to $|\Phi(Z, \varepsilon, 0)|^2$; $A(\varepsilon_1, \varepsilon_2)$ and $B(\varepsilon_1, \varepsilon_2)$ are rather complicated functions which depend on the invariants I_{ij}, J_{ij} and on the matrix elements

$$\langle \psi_f | \sum_{m,n} \frac{\tau_n^+ \tau_m^+}{r_{nm}} | \psi_i \rangle , \qquad r_{nm} = |\boldsymbol{x}_{nm}| ; \tag{2.10}$$

$$\langle \psi_f | \sum_{n,m} \tau_n^+ \tau_m^+ \frac{\boldsymbol{\sigma}_n \cdot \boldsymbol{\sigma}_m}{r_{nm}} | \psi_i \rangle . \tag{2.11}$$

The third possible matrix element

$$\langle \psi_f | \sum_{n,m} \tau_m^+ \tau_n^+ \frac{(\boldsymbol{\sigma}_n \cdot \boldsymbol{x}_{nm})\,(\boldsymbol{\sigma}_m \cdot \boldsymbol{x}_{nm})}{r_{nm}^3} | \psi_i \rangle ,$$

reduces to the second one due to the isotropy of the states $|\psi\rangle_{i,f}$. From the probability distribution the lifetime τ for no-neutrino double β-decay given by

$$\tau^{-1} = \int_0^{\varepsilon^0} \mathrm{d}\varepsilon_1 \int_0^{\varepsilon^0} \mathrm{d}\varepsilon_2 \, 2\pi \sin\theta \, \mathrm{d}\theta \, P(\boldsymbol{p}_1, \boldsymbol{p}_2)$$

where ε^0 is the decay energy release.

The evaluation of the matrix elements (2.10) and (2.11) can be done only in a very crude way [18, 25]. r_{nm} is usually approximated by the nuclear radius $R = 1.2 \cdot 10^{-13} A^{\frac{1}{3}}$; a more reasonable approximation might be to replace r_{mn} with $\langle r_{mn} \rangle$, the average distance of two neutrons in the initial nucleus ($\langle r_{nm} \rangle = 1/35\ \mathrm{MeV}^{-1}$). Further

$$\langle \psi_f | \sum_{m,n} \tau_m^+ \tau_n^+ | \psi_i \rangle = \sum_{\varrho} \langle \psi_f | \sum_m \tau_m^+ | \psi_\varrho \rangle \langle \psi_\varrho | \sum_n \tau_n^+ | \psi_i \rangle \simeq$$

$$\simeq \sum_{\varrho} |\langle \psi_\varrho | \sum_n \tau_n^+ | \psi_i \rangle|^2 \sim N_\varrho \, |\langle \psi_\varrho | \sum_n \tau_n^+ | \psi_i \rangle|^2 ,$$

and similarly for (2.11), N_ϱ being the number of states of the intermediate nucleus with appreciable single β-decay. From the corresponding matrix elements in single β-decay one estimates $|\langle \psi_\varrho | \sum_n \tau_n^+ | \psi_i \rangle|^2 \sim 10^{-2}$ and, assuming $N_\varrho = 10$, one gets

$$\left| \langle \psi_f | \sum_{n,m} \frac{\tau_n^+ \tau_m^+}{r_{nm}} | \psi_i \rangle \right|^2 \simeq 10^{-2} R^{-2} \simeq \left| \langle \psi_f | \sum_{n,m} \tau_n^+ \tau_m^+ \frac{\boldsymbol{\sigma}_n \cdot \boldsymbol{\sigma}_m}{r_{nm}} | \psi_i \rangle \right|^2 ,$$

where $|\psi_f\rangle$ and $|\psi_i\rangle$ are supposed not to belong to different eigenstates of T; otherwise $\langle\psi_f|\sum_{m,n}\tau_n^+\tau_m^+|\psi_i\rangle \equiv 0$.

These approximations are very crude: they might be easily wrong by a factor of 10^2 or more. The quantitative results that are obtained through them should therefore not be trusted too much. Below are given, as a function of A and Z, the lifetimes to be expected for a no-neutrino transition in this approximation in the two particular cases of

$$f_{\mathrm{I},s} = f'_{\mathrm{I},s} = +f'_{\mathrm{II},s} = -f_{\mathrm{II},s} = f_s; \quad f_{\mathrm{I},v} = -f'_{\mathrm{I},v} = -f_{\mathrm{II},v} = f'_{\mathrm{II},v} = f_v\,,$$

all other $f_i = 0$ and of

$$f_{\mathrm{I},v} = f'_{\mathrm{I},v} = -f_{\mathrm{II},v} = f'_{\mathrm{II},v} = f_v; \qquad f_{\mathrm{I},s} = \alpha^{-1}f'_{\mathrm{I},s} = -f_{\mathrm{II},s} = \alpha^{-1}f'_{\mathrm{II},s} = f_s\,,$$

(due to the symmetry of the problem no interference can take place between a Gamow-Teller and a Fermi matrix element).

$$(2.12)\qquad \tau \sim 4\cdot 10^{16\pm 2}\left(\frac{60}{Z}\right)^2\left[1-\exp\left[\frac{-2\pi Z}{137}\right]\right]\left(\frac{A}{150}\right)^{\frac{2}{3}}\left(\frac{8}{\varepsilon^0}\right)^5\left[\beta\left(\frac{1-\alpha}{2}\right)^2 f_v^2 f_s^2 + \right.$$
$$\left. + \left(\frac{1-\alpha}{2}\right)\left(\frac{1-\alpha^2}{2}\right) f_v f_s^3 + \gamma\left(\frac{1-\alpha^2}{2}\right)^2 f_s^4\right]^{-1} \text{ years};$$
$$\text{units } m_e = \hbar = c = 1;\ \beta \simeq 10,\ \gamma \sim \tfrac{1}{2}.$$

The first case is obtained if $\alpha = -1$; notice that the choice $\alpha = +1$, corresponding to a two-component theory with conservation of neutrinos gives $\tau^{-1} = 0$, as expected.

Let us now see briefly the corresponding calculations and approximations in the case of a double β-decay with neutrinos. In this case the matrix element is given, omitting symmetrization factors, by

$$(2.13)\qquad \langle\beta^-_{p_1\mu_1}\beta^-_{p_2\mu_2}\nu_{\delta_1 K_1}\nu_{\delta_2 K_2}(A, Z+2)\,|\,(A, Z)\rangle = \sum_{\varrho,i,j} R^{ij}(\varrho)\cdot(u_e^{\mu_1 +}(\boldsymbol{p}_1)O^i(f_{\mathrm{I},i} + f'_{\mathrm{I},i}\gamma_5)\cdot$$
$$\cdot Cu^{\delta_1 +}(\boldsymbol{K}_1))\,(u_e^{\mu_2 +}(\boldsymbol{p}_2)O^j(f_{\mathrm{I},j} + f'_{\mathrm{I},j}\gamma_5)Cu^{\delta_2 +}(\boldsymbol{K}_2))\,(W_\varrho - W_i + \varepsilon_1 + 1 - |\boldsymbol{K}|)^{-1}\,,$$

where

$$R_{ij}(\varrho) = \langle\psi_f|\sum_n \tau_n^+ \Phi^*(Z, \varepsilon_2, \boldsymbol{p}_2\cdot\boldsymbol{x}_n)\exp\left[-i(\boldsymbol{K}_2\cdot\boldsymbol{x}_n)\right]|\psi_\varrho\rangle\cdot$$
$$\cdot\langle\psi_\varrho|\sum_m \tau_m^+ \Phi^*(Z, \varepsilon, \boldsymbol{p}_1\cdot\boldsymbol{x}_m)\exp\left[-i\boldsymbol{K}\cdot\boldsymbol{x}_m\right]|\psi_i\rangle\,.$$

For the sake of simplicity, and without loss of generality, at least as far as the order of magnitude of the lifetime is concerned, we take $f_{\mathrm{IV}} = f_{\mathrm{I}a} = f'_{\mathrm{IV}} = f'_{\mathrm{I}a} = f$,

all other coefficients equal to zero [18]. We again expand the electrons' and neutrons' wave functions according to (2.3) and keep only the zero order term (here $p_{\max}\langle r\rangle_{\text{av}} \ll 1$ for both electrons and neutrinos); the states $|\psi_i\rangle$ and $|\psi_f\rangle$ will be again assumed to correspond to spin zero and to the same parity. Now, in the energy, denominator $|\boldsymbol{K}_i|$ no longer predominates; we shall assume, instead, that only low-lying intermediate states contribute substantially. $W_l - W_i$ will therefore be given a constant value and the sum of an intermediate state, carried out as before by closure. This approximation is probably worse than the corresponding one introduced in the no-neutrino case, but anyhow should not change too badly the order of magnitude of the result. In this way the transition amplitude is given in terms of quantities A, B [18] which contain the matrix elements

$$\langle\psi_f|\sum_{m,n}\tau_n^+\tau_m^+|\psi_i\rangle \tag{2.14}$$

and

$$\langle\psi_f|\sum_{n,m}\tau_n^+\tau_m^+\frac{\boldsymbol{\sigma}_n\cdot\boldsymbol{\sigma}_m}{r_{nm}}|\psi_i\rangle\,. \tag{2.15}$$

These expressions are approximated, as before, to 10^{-2}. Taking the square moduli, multiplying by the density in phase space of the final states and summing over all spin and momenta of the outgoing electrons and neutrinos one gets the lifetime, with a further, presumably small, simplification in the denominators to make the integrals simpler to evaluate.

$$\tau \simeq 1\cdot 10^{20\pm 2}\left(\mu_{\text{B}}-\mu_{\Lambda}+\frac{\varepsilon^0}{2}+1\right)^2\left(\frac{6}{Z}\right)^2\left(1-\exp\left[\frac{-2\pi Z}{137}\right]\right)\left(\frac{8}{\varepsilon^0}\right)^{10}\ \text{years}\,, \tag{2.16}$$

(in units $m_e c^2 = 1$)

the error $10^{\pm 2}$ coming expecially from the approximation in the nuclear matrix elements (2.14), (2.15). The dependence on ε^0 is again only approximatively valid.

It is easily seen, comparing (2.12) and (2.16), that the ratio between the two lifetimes is of the order of 10^{-5} for intermediate values of ε^0, Z and comparable nuclear matrix elements. It should also be stressed that, although the estimates of the two lifetimes given by formulae (2.12) and (2.16) might be very unaccurate, their ratio R has better chances to be close to reality, since the main source of error is the approximation of the two matrix elements (2.14) and (2.15) which appear quadratically both in the denominator and the numerator of R. Unfortunately what has to be compared with the experimental data are the lifetimes themselves, not their ratio, and this makes the interpretation of the experimental lower limits available extremely inconclusive.

For the sake of completeness, we quote the lifetimes to be expected in those double processes in which one, or two electrons are captured (units are chosen so that $m_e = c = \hbar = 1$):

a) $A + e^- \to (B + \nu) \xrightarrow[\text{virtual}]{} C + e^- (+\nu + \bar{\nu})$

1) no neutrino emission

$$\tau \simeq 8.6 \cdot 10^{26 \pm 2} Z^{-4} A^{\frac{2}{3}} \exp\left[\frac{2\pi(Z-60)}{137}\right] \varepsilon_0^{-4} \text{ years};$$

2) two neutrinos emitted

$$\tau \simeq 6 \cdot 10^{33 \pm 2} \left(\mu_B - \mu_A + \frac{1}{2}\varepsilon^0\right)^2 Z^{-1} \varepsilon_0^{\,7} \exp\left[\frac{2\pi(Z-60)}{137}\right] \text{years};$$

b) $A + e^- + e^- \to (B + e^- + \nu) \xrightarrow[\text{virtual}]{} \begin{cases} C + \gamma \\ C + \nu + \bar{\nu} \end{cases}$

1) no neutrino emission $\tau \simeq 3.6 \cdot 10^{36 \pm 2} Z^{-8} A^{\frac{2}{3}} \varepsilon_0^{-1}$ years

2) two neutrinos emitted $\tau \simeq 3 \cdot 10^{28 \pm 2} (\mu_A - \mu_B + \frac{1}{2}\varepsilon_0 - 1)^2 Z^{-6} \varepsilon_0^{-5}$ years.

Notice also that, within the approximation introduced, the lifetime for emission of two positive electrons is given by the same formula as the lifetime emission of negative electrons, both in the two-neutrino and in the no-neutrino case, sustituting Z with $-Z$.

We would like to close this section with a couple of remarks:

The estimates of the lifetimes for double β-decay as given in eq. (2.12) and (2.16) have been obtained under the assumption that both virtual intermediate transition are allowed. If this were not the case, the estimates should be increased by a factor $10^4 \div 10^5$ for each degree of forbiddingness; we shall comment briefly in Sect. **4** on the actual situation for some of the most promising triplets of isotopes.

The estimates (2.12) and (2.16) correspond to the extreme cases of a maximal non-conservation of the lepton number and of conservation of neutrinos respectively. We recall here again that the latter situation is obtained, *e.g.*, in the two-component theory with V-A interaction. These extreme cases correspond to a particular choice of the parameters f in the expression (1.1) for the Hamiltonian density. It is worth emphasizing that, with a proper choice of these parameters one can obtain, as an extimate for the lifetime for double β-decay, *any* value between those given by (2.12) and (2.16). This can be most easily seen if one lets, in eq. (2.12), the parameter vary continuously from -1 to $+1$, and remembers that the double β-decays with and without neutrinos are competing processes.

3. – Experimental results on the double β-decay.

Several experiments have been carried out during the last ten years in order to investigate the double β radioactivity of the more suitable isotopes. Some experiments lead to apparently positive results; but the conclusions so obtained were found to be inconsistent with the results of later experiments. Minimum half-lives for the various transitions were determined on the basis of negative results, but in only a few experiments can the minimum half-lives so obtained be compared with the theoretical estimates.

To check the theoretical predictions one should detect very weak activities, and this can be achieved only by increasing considerably the amount of material examined and/or by drastically reducing the background due to natural radioactivity and cosmic radiation. The difficulties are obviously connected with the techniques used in the experiment. For this reason it seemed worthwhile to us to analyze the experimental results so far obtained and to consider their possible improvement, taking into account the more important features of the various experimental techniques. We will consider separately the experiment based on the use of electronic counters, cloud chambers, nuclear emulsions and chemical separations.

3·1. *Counter experiments.* – The first experimental search, with counters, for the double β-decay was carried out by E. L. FIREMAN [26-28] on the possible transition $^{124}_{50}\mathrm{Sn} \rightarrow {}^{124}_{52}\mathrm{Te}$. Two samples were used in this experiment; a « working » sample enriched to 54% ^{124}Sn, and a « control » sample containing 0.4% ^{124}Sn for background measurements. Each of the samples was placed between a pair of thin mica window G.M. counters; the specimens were interchanged periodically. The two pairs of measurement counters were shielded by anticoincidence counters and by a Fe and Pb box. In a series of runs the enriched sample gave about 13% more coincidence counts than the depleted one; if this effect is interpreted as double β-decay of ^{124}Sn, an half-life of $(0.65 \pm 0.25) \cdot 10^{16}$ years is obtained. As we shall note later this apparently positive result is inconsistent with the results of several later experiments on the same transition. FIREMAN himself concluded later [46] that his own result was probably due to radioactive contamination of the enriched sample.

An experiment was performed by M. I. KALKSTEIN and W. F. LIBBY [29] on the same transition of ^{124}Sn in order to check FIREMAN's result. The experimental apparatus consisted of two thin walled G.M. counters operated in coincidence, surrounded by 11 guard counters in anticoincidence. The entire set-up was shielded with 20 cm of iron. Samples of enriched (95%) and of natural tin were placed alternately between the thin walled counters. The two measured activities were the same within statistical fluctuations; assuming

that the maximum activity is equal to twice the standard deviation one can calculate a minimum half-life between 1.7 and $2.4 \cdot 10^{17}$ years. An investigation was also made on the possibility of two single β transition through the intermediate isotope $^{124}_{51}$Sb. Antimony carrier, added to old stannic cloride, was separated and counted for the 60-day β activity of ^{124}Sb. By this experiment a minimum half-life of about 10^{17} years was set for the single β-decay of ^{124}Sn (see Fig. 12*e*).

Proportional counters have seldom been used in experiments on the double β-decay. In fact these detectors can be used to measure the energy of slow electrons, such as the photoelectrons for X-rays; therefore they are particularly suitable to experiments on double electron-capture or negatron-capture, positron-emission.

An experiment of this type was carried out by A. BERTHELOT *et al.* [30] on the double electron-capture (*KK*) $^{64}_{30}\mathrm{Zn} \rightarrow {}^{64}_{28}\mathrm{Ni}$. This transition should be accompanied by the emission of an 8 keV X-ray from the filling of *K*-shell of ^{64}Ni. A brass tube, with the internal surface covered by a thin layer of natural zinc (48.89% ^{64}Zn) was placed inside a gas proportional counter. The pulses from the counter were analysed on a 10-channel pulse height analyser. The proportional counter was surrounded by a layer of guard G.M. counters operated in anticoincidence; the entire set-up was shielded by 20 cm of lead. To calibrate the counter the counting rate of an internal calibration source of ^{65}Zn was compared with the counting rate due to an internal source of variable intensity. The comparison showed that an internal activity could be detected only if it exceeded the background by at least 12%. No detectable activity of ^{64}Zn was discovered above the background, which was determined by the use of a brass tube without ^{64}Zn; a lower limit of $8 \cdot 10^{17}$ years was set for the half-life of the double electron-capture of ^{64}Zn. A. BERTHELOT *et al.* considered in addition the competitive possibility of positive double β-decay of ^{64}Zn; this possibility can however be ruled out on the basis of recent measurements of the masses of ^{64}Zn and ^{64}Ni (see Sect. **4**).

H. SELIG [31-32] used a proportional counter similar to that of A. BERTHELOT *et al.* to investigate the total activity of many elements. His results can be used also in the searchs for double β-decays, because they can often rule out the possibility of the sequence of two single β-decays through the intermediate isotope. As in BERTHELOT's experiment, a thin layer of the element to be examined was deposited on the internal surface of a tube placed inside the proportional counter. The counter was shielded by 15 G.M. counters in anticoincidence and layers of iron (5 cm), lead (7.6 cm) and mercury (1.3 cm). Pulses from the proportional counter were measured by means of a 48-channel pulse height analyser; energy measurements were made only on X-ray photoelectrons; no attempt was made to measure the energy of decay electrons.

Measurements were made on natural calcium carbonate (43 grams) to

investigate the possible β-decay $^{48}_{20}Ca \rightarrow {}^{48}_{21}Sc$. It is worthwhile to note that this activity, if detected, would make extremely uncertain every positive result on double β-decay of ^{48}Ca obtained with « not coincident » methods such as nuclear emulsions and chemical separations (see 3.3 and 3.4). The spin-parity of the ground state of ^{48}Ca is $0+$ and that of ^{48}Sc is $6+$ or $7+$ [67-68]; the single β-decay of calcium should therefore be sixth forbidden (see Fig. 12*a*)). Selig found for calcium a very small excess activity above the background and set a lower limit of $2 \cdot 10^{14}$ years for the half-life of ^{48}Ca for every kind of decay, and of $2 \cdot 10^{16}$ years for the decay of ^{48}Ca to the 48-hour ^{48}Sc.

A negative result was obtained by Selig on the single β transition $^{96}_{40}Zr \rightarrow {}^{96}_{41}Nb$; direct counting of zirconium gave a lower limit of $2 \cdot 10^{16}$ years for the half-life. A similar negative result on the decay of $^{100}_{42}Mo$ to the 17-second $^{100}_{42}Tc$ gave a minimum half-life of $2 \cdot 10^{16}$ years. The result obtained by Selig on the transition $^{130}_{52}Te \rightarrow {}^{130}_{53}I$ can also be considered as negative because the total activity in excess on the background is less than the standard derivation; a minimum half-life of $4 \cdot 10^{16}$ years was obtained.

The first experiment which measured the sum of the kinetic energies of the two electrons was carried out by R. M. Pearce and E. K. Derby [33] with scintillation counters. The source, consisting of a thin foil of tin, enriched to 95% ^{124}Sn, was placed between two scintillation counters employing anthracene crystals. The counters, which were operated in coincidence, were shielded from above with 16 cm lead and in other directions whit 8 cm of lead; the output was recorded on a 18-channel analyser. The background was evaluated by replacing the tin foil with an aluminum foil of the same thickness. The spectrum of the pulse from the two counters, added electronically, did not show a peak at an energy corresponding to the total energy release. In order to obtain a lower limit for the half-life the geometrical efficiency of the counters and the trasmission factor of the foil were evaluated; a minimum half-life of $(0.45 \pm 0.15) \cdot 10^{17}$ years was obtained in this way.

Results on the double β-decay of ^{124}Sn and ^{96}Zr were obtained by J. A. McCarthy using scintillation counter methods [34-36]. The experimental arrangement is shown in Fig. 2. The sample was placed between two stilbene crystals viewed by photomultiplier tubes whose output pulses were added electronically. Four an-

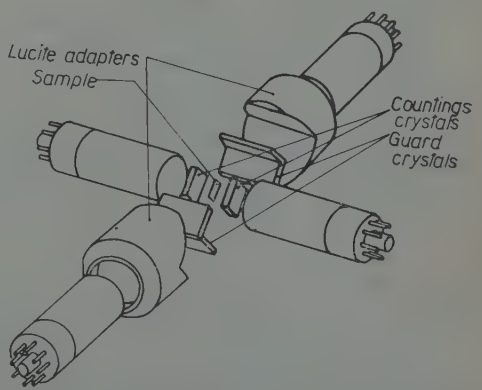


Fig. 2. – Schematic view of counting crystals and anticoincidence counters used by Mc-Carthy [34-36].

ticoincidence scintillation counters surrounded the source and the « measurement » crystals (only two guard counters are shown in Fig. 2). The equipment was shielded with 0.6 cm of steel and 10 cm of lead; a 0.16 cadmium sheet was placed inside the shield to absorb thermal neutrons. Coincident pulses were analysed by means of a 30-channel pulse height analyser. The total activity was also measured, without regards to coincidence.

In a first run McCarthy compared the activity of a sample of tin enriched to 95% ^{124}Sn and a sample of natural tin. The two specimen were changed alternately by moving one of the scintillation counters. No significant difference between the two coincidence activities was observed in the region of energy corresponding to the energy release of the transition. The value for energy release used by McCarthy was 1.5 ± 0.4 MeV; if we use the more recent value 2.31 ± 0.39 MeV we observe no statistically significant difference between the two counting rates. On the basis of this negative result one can set a lower limit of $1.5 \cdot 10^{17}$ years for the half-life of ^{124}Sn.

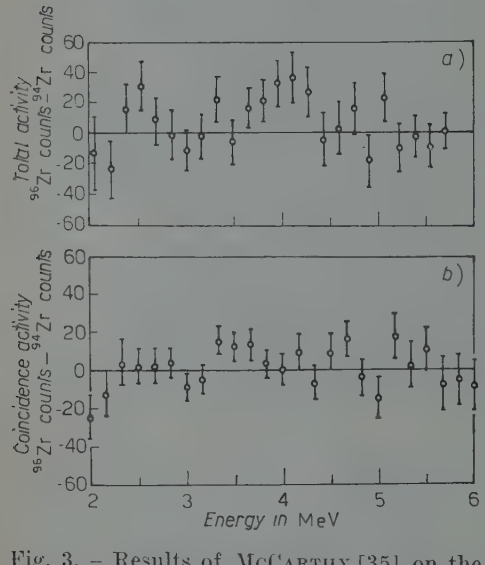


Fig. 3. – Results of McCarthy [35] on the transition $^{96}_{40}$Zr - - $^{96}_{42}$Mo. Differences between total and coincidence activities with ^{93}Zr and ^{91}Zr samples.

In a second run McCarthy investigated the total and coincidence activity of a zirconium sample enriched to 89.5% ^{96}Zr and of a zirconium sample enriched to 97.9% ^{94}Zr, as control sample. The differences between coincidence and total activity of the « working » and « control » sample are shown in Fig. 3. The apparent peak between 3.3 and 4.3 MeV in the coincidence spectrum seems to indicate a neutrinoless double β-decay (the energy release is (3.4 ± 0.3) MeV). This peak cannot be explained as a statistical effect; it can however be ascribed to an internal conversion of 3.8 MeV γ-rays or to internal pair production. The possibility of a radioactive contamination of the ^{96}Zr can be ruled out, because this effect should be detected also in the total activity. If the peak in Fig. 3 is assumed to be an experimental evidence for the double β-decay, a mininum half-life of $(6 \pm 2) \cdot 10^{16}$ years is obtained.

A later experiment on the double β-decay was carried out by McCarthy with an experimental set-up very similar to the earlier one. The equipment was shielded by 2.5 cm of steel and 7.5 cm of lead during the first part of

the experiment; the next part of the experiment was performed under a 174-meter building and the equipment was shielded with 0.4 cm of steel and 7.5 cm of lead. The scintillation counters were surrounded by an arrangement of G.M. tubes in anticoincidence. Double β-decay activity was investigated in a sample of $CaCO_3$ enriched to 84.3 % ^{48}Ca; background measurements were made on a similar sample depleted to 0.01 % ^{48}Ca. The counting difference between the two coincidence activities is shown in Fig. 4 as a function of the energy; the standard deviation is given for each point. The curve drawn in Fig. 4 indicates the distribution theoretically predicted by McCARTY for the decay of ^{48}Ca with a half-life of $1.6 \cdot 10^{17}$ years. The small peak at 3.1 MeV was calculated for a decay from the ground state of ^{48}Ca to the first excited level of ^{48}Ti (see Fig. 12 *a*)) with the assumption that the nuclear matrix element for this transition is the same as that for the ground state transition. The peak corresponding to a transition between the ground states was observed at an energy of (4.1 ± 0.3) MeV which is in good agreement with the energy release calculated on the basis of the mass spectroscopic data (Sect. 4). The agreement between the theoretical calculation and the experimental results seems quite good; McCARTY's positive result can be considered as the only possible evidence obtained so far for double the β-decay without neutrino emission. The probability that the difference in activity of Fig. 4 was due to statistical fluctuations is less than 10^{-6}. The only possibility other than double β-decay seems to be a contamination of the ^{48}Ca sample or the back-ground due to cosmic radiation.

An experiment was carried out by M. AWSCHALOM [37] to check the positive results obtained by McCARTHY on the $\beta^-\beta^-$ activities of ^{96}Zr and ^{48}Ca. An

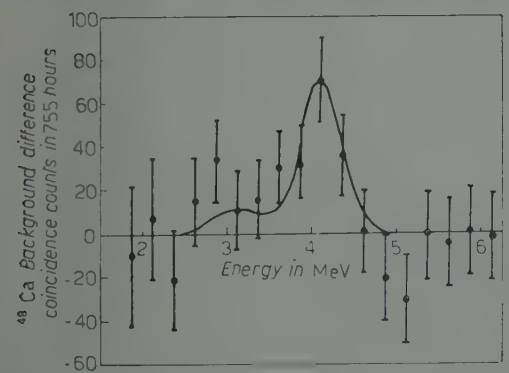


Fig. 4. – Difference between the coincidence activities of a ^{48}Ca and a ^{44}Ca sample. From McCARTHY [36].

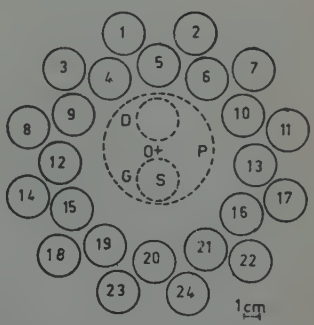


Fig. 5. – Schematic view of the apparatus of DETOEUF and MOCH [38].

experimental apparatus, very similar to that used by McCARTHY, was placed in a mine at a depth of 900 m w.e. Thus the background due to the cosmic radiation

was almost completely eliminated; the low energy γ-ray activity due to ^{40}K existing in the mine was strongly reduced by an iron shield of 2.5 cm. The measurements were carried out on three specimens: *a*) a natural calcium sample, *b*) a calcium sample enriched to 85.28% ^{48}Ca, *c*) a zirconium sample enriched to 85.25% ^{96}Zr. The AWSCHALOM's results on the coincidence activities of the tree samples in the energy regions ((2.5 ÷ 4.5) MeV) and ((3 ÷ 4.75) MeV) are listed in Table I. No significant activity above the background was found; the lower limits for the half-life were calculated from a net excess activity of the sample over the inactive one of twice the standard deviation. The discrepancy between MCCARTY's and AWSCHALOM's results can be explained if one assumes that most of the background is due to cosmic radiation, since this radiation is extremely reduced in the AWSCHALOM's experiment. This explanation is fully consistent with the later results of E. I. DOBROKHOTOV (see pag. 159).

TABLE I.

Sample	Running time (h)	Detection efficiency (%)	Counting rate (events/100 h)		Half-life (years)
			(2.5 ÷ 4.5) MeV	(3.0 ÷ 4.75) MeV	
Total activity					
^{48}Ca	303.97	81	24.0	12.8	$> 1.1 \cdot 10^{18}$
nat. Ca	207.27	—	26.5	14.2	—
^{93}Zr	319.24	60	25.5	13.4	$> 0.36 \cdot 10^{18}$
Coincidence activity					
^{48}Ca	640.49	35	7.4	4.4	$> 2 \cdot 10^{18}$
nat. Ca	252.13	—	4.8	4.4	—
^{96}Zr	503.52	25	5.5	4.9	$> 0.5 \cdot 10^{18}$

The double β-decay $^{116}_{48}Cd \rightarrow {}^{116}_{50}Sn$ was investigated by J. F. DATOEUF and R. MOCH [38]. Coincidence activities were measured (Fig. 5) by means of a pair of vertical scintillation counters with axis S, viewing each side of a disk P which interchanged the two samples D and G. The scintillation counters were surrounded by a double layer of G. M. counters in anticoincidence; the whole set-up was shielded with 10 cm of lead. The samples consisted of two thin disks of cadmium enriched to 81% ^{116}Cd and to 95.3% ^{114}Cd, respectively. The thickness of the scintillation crystals was determined so that the sums of the energy losses in the two scintillators, due to minimum ionizing rays, were about 5.5 MeV. Thus the background of coincidences due to cosmic rays was outside the region of energies to be investigated.

The heights of coincident pulses from the scintillation counters were sent

to the orthogonal plates of an oscilloscope and recorded on a slowly moving film. The excess activity of the active sample over the inactive one was less than the standard statistical fluctuation. DATOEUF assumed, as upper limit for a detectable activity, the activity which would give the observed excess with a probability 10 times less than the probability of a corresponding excess due to statistical fluctuation. In this rather complicated way a minimum half-life of $3 \cdot 10^{16}$ years was obtained.

An experiment on the double transition $^{150}_{60}\mathrm{Nd} \rightarrow {}^{150}_{62}\mathrm{Sm}$ was carried out by C. L. COWAN *et al.* [39-40]. A sample of neodimium enriched to 82.23% ^{150}Nd and a control sample of natural neodimium depleted to 0.065% ^{150}Nd were used for $\beta^-\beta^-$ and background measurements, respectively. The two specimens were immersed in a liquid scintillator and viewed on each side by two pairs of photomultipliers; the samples were periodically interchanged during the measurements. The whole apparatus was immersed in a 600-liter tank of liquid scintillator viewed by 12 guard photomultipliers operated in anticoincidence. The tank was shielded with 10 cm of lead and bismuth. Coincident pulses from the measurement counters were added electronically and analysed by means of a 100-channel pulse height analyser if the sum of their energies was between 0.35 and 8 MeV. The coincidence resolving time was 0.5 μs. The number of events is plotted as a function of the sums of energies in Fig. 6. The peak at about 5.6 MeV was attributed by COWAN to two successive Compton scatterings of the 7.6 MeV γ-rays arising from neutron capture in iron. On the basis of the spectroscopic values of the masses available at the time the experiment was performed, a peak should be detected at an energy of (4.4 ± 0.8) MeV for a neutrinoless decay. According to more recent mass measurements (see Sect. 4) the total decay energy should be (3.65 ± 0.09) MeV. The distribution curve is flat between 3.5 and 5 MeV (Fig. 6); thus Cowan's result can be considered as completely negative. The authors calculated a maximum activity corresponding to the energy release in an interval equal to the experimental energy resolution in the measurement of the sums of energies. A minimum half-life of $2 \cdot 10^{18}$ years was obtained; this result is also valid for the more recent mass determinations. It is worthwhile to note, however, that COWAN did not make any comparison with background activity.

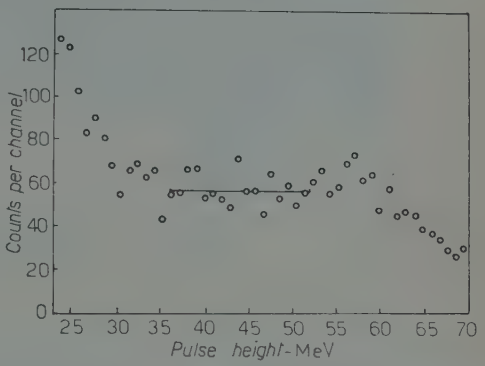


Fig. 6. – Distribution of the coincidence pulses recorded by COWAN *et al.* [39] as a function of the energy sum.

Thus the evaluated lower limit of the half-life for the transition is valid only with the assumption of a neutrinoless double β-decay. If a two-neutrino decay is assumed, the minimum half-life calculated by COWAN must be considered as overestimated.

The most recent and interesting experimental results on the double β-decay were obtained by E. I. DOBROKHOTOV *et al.* [41-42] with two elegant experiments on the transition $^{48}_{20}Ca \rightarrow {}^{48}_{22}Ti$. The specimens used in the first experiment were two « working » samples of 423 and 259 mg of Ca enriched to 76.2 % ^{48}Ca and two « control » samples of equivalent masses of calcium enriched to 99.9 % ^{44}Ca. Two runs were made at the same time: one on the earth's surface under a shield of 15 cm of lead and 3 cm of steel and the other underground at a depth of 33 m. In both runs, the working and control activities were measured with pairs of scintillation counters 1 and 2 viewing the sample on each side (Fig. 7). The samples were interchanged between the scintillators by the motor 3.

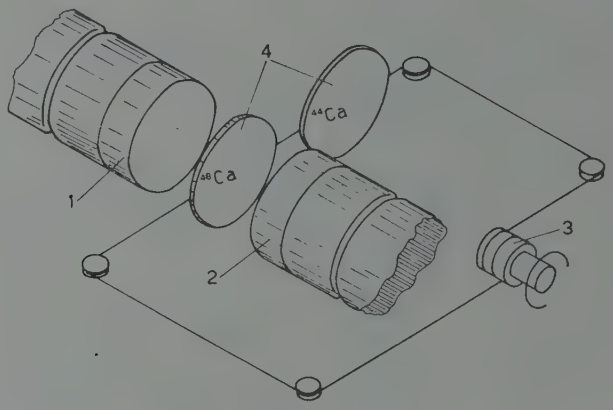


Fig. 7. – Sketch of arrangement of the measuring head. From DOBROKHOTOV *et al.* [42]: 1) 2) scintillation counters; 3) shaft of motor for shifting the specimens; 4) specimens.

In the surface experiment the apparatus was immersed in a 3-liter tank of liquid scintillator, the tank as a whole being viewed by two photomultipliers in anticoincidence with the measurement counters. In the underground experiment a tank containing 65 liters of liquid scintillator, scanned by 21 photomultipliers, was used (Fig. 8). The pulses from the two measurement photomultipliers, which were in coincidence within 0.5 μs and in anticoincidence with the guard counters, were sent to a double oscilloscope and recorded on film. The activities of the working and control sample were compared, taking into account the resolution of the apparatus, in a region of sums of energies between 3.2 and 4.8 MeV. No significant excess activity of ^{48}Ca sample was found above the statistical fluctuations. Furthermore the distribution of

coincident counts as a function of the sum of energies showed the random character of the coincident events for both samples. A lower limit of $1 \cdot 10^{18}$ years for the half-life was set, assuming, as usual, a maximum excess activity of twice the standard deviation.

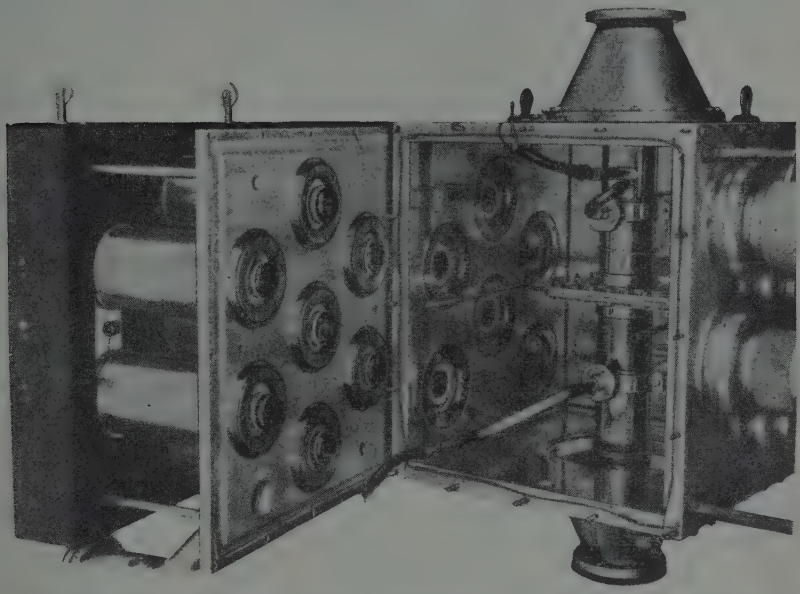


Fig. 8. – Measuring and guard photomultipliers in the Dobrokhotov's experiment. (Courtesy of Dr. E. I. Dobrokhotov).

A later experiment was carried out by the same authors with an apparatus very similar to that used in the previous underground experiment. Great care was taken to study and to decrease the background counting and to improve the energy resolution of the measurements. Two measuring counters very similar to those used in the previous experiment were immersed in a 70-liter tank of liquid scintillator which was wiewed by 21 guard photomultipliers operated at the same amplification in anticoincidence. The equipment was placed underground at a depth of 65 m w.e. and shielded with a layer of 10 cm of lead. The coincidence resolving time between the two measurement counters was reduced to 0.2 μs. The energy calibration of the measurement counters was performed with a sample of ^{137}Cs, by means of the monocromatic electrons of ^{137}Ba (0.625 MeV). A second calibration point was obtained at 3.5 MeV, by means of cosmic ray muons, crossing both scintillators operated vertically, with a minimum ionization and a consequentely costant energy loss. A half-value width of 11% was obtained for 4 MeV electron, by extrapolation. The linearity of the measurement counters was checked with a β-ray spectrometer, operated with a source of ^{137}Cs.

A series of control runs was carried out in order to check the electron counting efficiency, the calibration of the scintillation counters and the shielding effectiveness:

a) the electron efficiency of the measuring scintillation counters was checked by comparison with the efficiency of thin-walled G.M. counters. The efficiency was found to be 100%.

b) The whole apparatus was fully actuated and a photographic registration of the energy spectrum of each counter was taken on the surface of the earth. The two spectra as well as the sums of the amplitude values of the pulses from each channel were compared: the agreement between both the spectra and the totals indicated that the two measurement counters were calibrated identically.

c) Measurements were made on the shielding effectiveness of the anti-coincidence counters and of the screens. It was demonstrated that the background of the coincident pulses was lowered by a factor of 23 by placing the apparatus underground. The lead shield led to a further reduction by a factor 5 and the anticoincidence counters by a factor of 3.3.

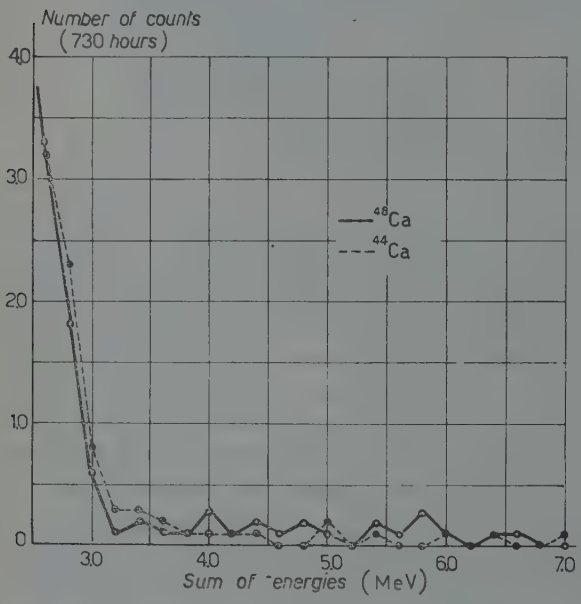


Fig. 9. – Results of DOBROKHOTOV *et al.* [42] on the coincidence activities of ^{48}Ca abd ^{44}Ca samples.

The same 423 mg specimens of the preceeding experiment were used as « control » and « working » samples. The spectra obtained by DOBROKHOTOV in 730 h of measurement are shown in Fig. 9. To interpret these results the authors started from the assumption of a neutrinoless double β-decay of ^{48}Ca. The total energy release was assumed to be (4.3 ± 0.1) MeV on the basis of mass spectroscopic data [57]. The mean energy loss of the two emitted electrons in the specimen was calculated to be 0.4 MeV, assuming no angular correlation in the emission. Although the calculated energy resolution for the added pulses from the measurement counters was 0.35 MeV, a figure of 0.5 MeV was used in the calculations

to provide a margin of error. The spectra of Fig. 9 was analysed in the energy interval between 3 and 4.4 MeV. In this energy region 11 events were registered with the specimen enriched in ^{48}Ca and 12 events with the control specimen anriched in ^{44}Ca. The difference between measurement and control coincidence counting was consequentely $(-1.4 \pm 6.6) \cdot 10^{-3}$ count per hour.

The minimum half-life was evaluated from the formula

$$\tau = \ln 2 \frac{N_0 k m \eta}{A \Delta n}, \tag{3.1}$$

where m is the mass of the sample (in grams), Δn is the difference in counts between the measurement and the control samples for unit time, N_0 is Avogadro's number, A is the mass number, k is a factor depending on the geometry of the counters; the coefficient η depends on the « transparency » of the specimen to decay electrons. Both k and η were calculated from an isotropic angular distribution of the electrons; the energy distribution between the two electrons calculated by YA. B. ZEL'DOVICH *et al.* [43] was used to calculate η. If one assumes, as usual, a maximum activity equal to twice the standard error, a minimum half-life of $7 \cdot 10^{18}$ years is obtained from (3.1). If however the angular distribution calculated by H. PRIMAKOFF and S. P. ROSEN is used, the minimum value of the half-life obtained in this and other experiments should be lowered by a factor of about 3.

It seems to us worth to emphasize that DOBROKHOTOV's result on the minimum half-life was obtained on the assumption of a neutrinoless double β-decay. If we start with the assumption of the emission of two neutrinos we cannot confine our considerations to the region of the spectrum corresponding to the total energy release (Fig. 1). The comparison between measurement and control activities must be made in a wider range of energy. A lower value for the minimum half-life would result.

* * *

From the analysis of the various counter experiments one can see that only in the first searchs G.M. counters were used to detect double β a ctivities In fact scintillation counters are much more suitable for a search of this type because they can detect the energies of the decay electrons. Proportiona counters can be used only to investigate double decays of particular type, such as double electron-captures. One of more important features of the experiments with the scintillation counters is the detection of the « time coincidence » in the emission of the two electrons. The check of this coincidence can be accomplished electronically within very short time intervals (0.2 μs in the last experiment of DOBROKHOTOV); a very sharp reduction of the background can be obtained in this way. Better results could be reached

with the recent improvements of the electronic technique which permit a further reduction in the coincidence resolving time [44]. It should be noted however that in the majority of counter experiments a further reduction of the background cannot bring a conclusive improvement to the results. In fact the limitation on the maximum detectable double activity is mostly due to the rather small quantity of specimen that can be investigared. For instance, in the last experiment of E. I. DOBROKHOTOV *et al.* the background is so sharply reduced that the limitation of the value of the minimum half-life can be ascribed only to the value of mass m in eq. (3.1). It does not seem worthwhile in order to investigate more material to increase the thickness of the specimen, because in this way the resolution in the measurements of the sum of the energies would be decreased. The only possibility for improvement seems so to be an increase in the sizes of the scintillation crystals in order to scan more material. The results could probably be considerably improved by using the double active sample itself as scintillator. A crystal scintillator of ^{48}Ca or a gas scintillator of ^{136}Xe could be used for experiments of this type.

The most relevant limitation in the experiments with electronic counters is the difficulty of localizing with good precision the point of origin of the two electron tracks. A better check of the « special coincidence » of the event can be achieved with the « visual » techniques as the cloud chamber or the nuclear emulsions. For this reason a comparison between the activity of the enriched sample seems to us always necessary in counter experiments.

A very conclusive reduction of background events could be reached by using magnetic devices in order to check the charge identity of the two decay electrons. In fact the majority of background is often due to negatron-positron pairs. An improvement of the experiments with counters in this direction is, in our opinion, of considerable interest.

3·2. *Cloud chamber experiments.* – The first experiment with a cloud chamber was carried out by J. S. LAWSON [45] who searched for the double β-decay of a sample of tin enriched to 83 % ^{124}Sn. In about 9 000 pictures, corresponding to a total effective sensitive time of 4 000 s, 39 possible double β-decay events were found. All these events were rejected for one of the following reasons (very interesting from the point of view of background reduction): *a*) the two tracks had a different origin; *b*) the two tracks were simulated by an electron circling twice through the sample; *c*) the tracks were of different age; *d*) the two tracks were split from one by the clearing field. From the Lawson's negative result a minimum half-life between 10^{16} and 10^{17} years for the double decay of ^{124}Sn can be estimated.

E. L. FIREMAN and D. SCHWARTZER [46] carried out a cloud chamber experiment in order to check the earlier positive result obtained by FIREMAN

with G.M. counters on the transition $^{124}_{50}Sn \rightarrow {}^{124}_{52}Te$. A cloud chamber, operated with helium and ethyl alcohol in order to reduce the multiple scattering, was placed in a magnetic field of 200 gauss. Three samples were used: two natural tin sheets, one 0.0075 cm thick and the other 0.002 cm thick, and a sheet of tin 0.004 cm thick enriched to 95.04% ^{124}Sn. Different runs were made with the sample placed across the chamber and acting as an electrode of the clearing field. An open counter was initially placed on either side of the sample; later two banks of three thin-walled counters were used. The chamber was triggered by coincidence pulses from the two sets of counters; an effective measurement time equivalent to more than 1 500 times the sensitive time of a random expanded chamber was obtained in this way. Five events with two electron tracks from the same origin and with the same sign were detected; in two of these events, however, the curvature had the wrong sign for negatrons leaving the source. These events can be interpreted as due to electrons passing through the source and having the direction of the curvature changed by gas scattering after leaving the source. The remaining three events seem to be due to pairs of negatrons leaving the source from the same point, but one cannot rule out the possibility of background pairs, with the positron having the sign of its curvature changed by scattering. In fact the sums of the energies of the electrons seem very low ($\approx 1/10$ MeV) in comparison with the total energy release. If the three obtained events are interpreted as double β-decays, a half-life of $(2 \div 5) \cdot 10^{17}$ years can be estimated; this result contradicts the earlier result obtained by FIREMAN with G.M. counters (see pag. 151).

R. G. WINTER carried out two cloud chamber experiments on the double β-decay of various isotopes. In the first experiment [47] the transition $^{108}_{46}Pd \rightarrow {}^{108}_{48}Cd$ and $^{110}_{46}Pd \rightarrow {}^{110}_{48}Cd$ were studied. A sample of 248.5 g of palladium, containing 26.7% ^{108}Pd and 13.5% ^{110}Pd, was placed in a Wilson cloud chamber. Only one possible double β-decay was detected from the scanning of about 10 000 photographs, corresponding to about 1 400 s of effective measurement time. The authors noted that the probability that this event was simulated by a negatron-positron pair due to the γ-rays of the cosmic radiation was about 0.6. If the event is interpreted as a double β-decay, one can calculate a minimum half-life of $1.1 \cdot 10^{18}$ years for the decay of ^{108}Pd and of $6 \cdot 10^{17}$ years for the decay of ^{110}Pd. It is worth noting that WINTER's results have lost most of their interest after the more recent mass measurements on the isotopes of palladium and cadmium. In fact the energy release for the decay of ^{110}Pd on the basis of the more recent data (Sect. 4) is (1.48 ± 0.53) MeV instead of 2.5 MeV, as assumed by WINTER; the transition $^{108}_{46}Pd \rightarrow {}^{108}_{48}Cd$ is energetically forbidden because the mass of ^{108}Cd exceeds the mass of ^{108}Pd.

The second experiment [48] was an attempt to detect the double β activity of the isotopes 106 and 116 of cadmium and 92 and 100 of molybdenum. The following transitions were considered: *a*) $^{116}_{48}Cd \rightarrow {}^{116}_{50}Sn$ $(\beta^-\beta^-)$ with an energy

release ε_0 of (2.61 ± 0.44) MeV on the basis of the more recent measurements (value used by WINTER [58] $\varepsilon_0 = (2.6 \pm 0.2)$ MeV); *b)* $^{106}_{48}Cd \rightarrow {}^{106}_{46}Pd$ $(\beta^+\beta^+)$ with an energy release of (0.64 ± 0.46) MeV (WINTER: $\varepsilon_0 = (0.8 \pm 0.2)$ MeV); *c)* $^{100}_{42}Mo \rightarrow$ $\rightarrow {}^{100}_{44}Ru$ $(\beta^-\beta^-)$ with an energy release of (2.3 ± 0.2) MeV from WINTER (no later result available); *d)* $^{94}_{42}Mo \rightarrow {}^{92}_{40}Zr$ $(\beta^+\beta^+)$ with an energy release of $-(0.42 \pm$ $\pm 0.53)$ MeV (no value from WINTER's paper).

Two runs were made with a 24 cm diameter Wilson cloud chamber placed in a magnetic field of 790 gauss. In the first run three thin foils of natural cadmium containing 2.27 g of ^{116}Cd and 0.36 g of ^{106}Cd were stretched across the chamber. The chamber was not sensitive in the layer of gas immediately adjacent to the foils, because the supersaturation was lost in that layer earlier than in other region of the chamber. It was however possible to locate the point of origin of the two electron tracks with an error of less than 0.4 cm. 12 352 acceptable stereoscopic pictures, corresponding to an effective measurement time of about 4 450 s were taken. 24 events with the features of double β-decay and one possible positive double β-decay were found. Each event is represented in Fig. 10 as rectangle of unit area with a width proportional to the uncertainty of the energy measurement for that event. The solid line, obtained by adding the heights of the rectangles whenever they overlapped, gives the experimental distribution of the sums of the two electron energies as a function of the energy.

A very similar distribution was obtained in the second run made with three foils of natural molybdenum (4.5 g of ^{100}Mo and 7.14 g of ^{92}Mo) stretched across the chamber. 35 possible negative double β-decay events and no possible positive double β-decay event were found on 13 402 acceptable stereoscopic pictures, corresponding to an effective measurement time of 7 240 s.

For a correct interpretation of his results, WINTER carried out a detailed investigation on the background. He recorded the number of single electrons coming out from the foils, and calculated the number of events that can be simulated by random coincidences, negatron-negatron scattering, double Compton scattering or single Compton scattering followed by photoelectric emission, and change of curvature of one of the electrons of background pairs due to gas scattering. The distribution of the various contributions to the background in cadmium experiment as a function of the energy is shown in Fig. 11; the effect due to the gas scattering is not shown, because it is negligible. The sum of all contributions is reported as a function of the energy in Fig. 10; it lies above the experimental distribution on most of the spectrum. Thus it seems not unreasonable that the background was overestimated. Since no maximum was observed in the region of energy corresponding to the energy release, the result can be considered as completely negative. Starting from the hypotesis of a neutrinoless double β-decay a minimum half-life of $1 \cdot 10^{17}$ years for ^{116}Cd and of $3 \cdot 10^{17}$ years for ^{100}Mo can be evaluated from the number of

events with sums of energies corresponding to the energy release. In the same way lower limits of $6 \cdot 10^{16}$ years and $4 \cdot 10^{18}$ years were calculated for the positive double β-decay of ^{106}Cd and ^{92}Mo respectively.

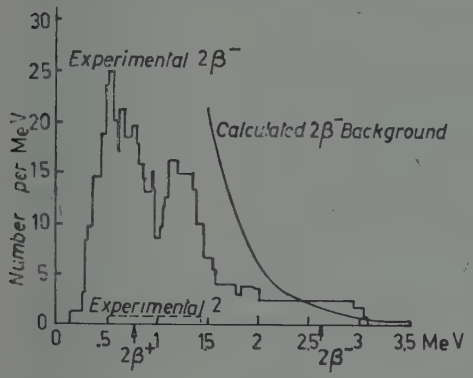


Fig. 10. – Esperimental results and calculated background in WINTER's experiment [48]. Transitions $^{116}_{48}Cd \rightarrow {}^{116}_{50}Sn$ and $^{106}_{48}Cd \rightarrow {}^{106}_{46}Pd$.

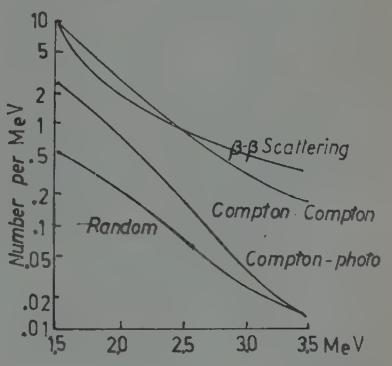


Fig. 11. Contributions to the double negatron background in WINTER's experiment.

* * *

From the consideration of the preceeding experiments one can easily see that one of the most favorable features of the cloud chamber experiments lies in the possibility of identifying the sign of the emitted particles. The majority of the background due to pairs from γ rays of the cosmic radiation or of the natural radioactivity can thus be eliminated. The technique can be improved by using low pressure chambers expanding in the vacuum, filled with a gas having a low atomic weight (hydrogen or elium) and by using methyl alcohol instead of ethyl alcohol. In this way the multiple scattering is strongly reduced and it should be possible to improve the measurements of the momentum of the electrons and to reduce the contribution of the background due to change in the curvature of the tracks.

With a cloud chamber in good working condition, it is easy to identify the « spatial coincidence » of the points of origin of the two electron tracks with a good precision (of the order of 1 mm). The « time coincidence » of the event can be checked by comparing the « age » of the two tracks; a precision of a few hundredths of a second can be obtained. For this reason the cloud chamber can be considered a quite suitable instrument for background reduction. The more relevant limitation in the use of this detector for the investigation of weak activities, and in particular of the double β-decay, arises from the sensitivity time which is always very short in comparison with the recicling time. Lower limits on the half-life could be reached by using large size multiplate cloud chambers containing large amounts of double β-active

samples, and by reducing the recicling time (quick compression, overcompression, gas exchange etc.). Needless to say, a chamber triggered on coincidence pulses corresponding to double β-decays, instead of a random expanded one, can be used; this seems however rather difficult to achieve if one wants to use considerable amounts of double β-active sample.

A detector that seems to us particulary suitable for double β-decay experiments is the continuosly sensitive diffusion chamber. We would like to suggest an experiment on the double β-decay to be carried out underground and under a suitable shielding with a diffusion cloud chamber filled with a double β-active gas (^{136}Xe for instance) placed in a magnetic field. With this experiment the same material would be used as radioactive sample and as detector. The « spatial coincidence » of the possible double β-decay events could be verified within few tenths of millimeter, with a very effective reduction of the background. The greatest difficulties in this type of experiment would arise from the problem of shielding and from the tremendous amount of pictures to be scanned.

3·3 *Nuclear emulsions experiments.* – The only experiment on the double β-decay with nuclear emulsions was carried out by J. H. FREMLIN and M. C. WALTERS [49]. The technique employed was to embed a few grams of ββ-active sample into recesses in graphite blocks facing the emulsion layer during the exposition. With this technique was possible to avoid direct chemical contact of the sample with the emulsion and to accomplish long exposures (of a few months duration) without loss of photographic sensitivity. The plates were placed in a mine at a depth of about 570 m and were shielded with layers of steel and lead. In this way it was possible to obtain a tremendous background reduction: the number of electron tracks was reduced from about 200 to 2.5 tracks per mm^2 of a 200 μ thick emulsion per day.

The plates were developed after exposures of duration ranging from two to eight weeks and scanned for the electrons crossing the upper surface of the emulsion layers. Electron tracks both in the regions exposed to the samples and in those which had faced the pure grafite, were counted. Exposures of different duration were made for each sample in order to detect the possible contamination by shortlived intermediate radio-elements. Contamination by uranium, thorium and their products were also investigated on the basis of their α-radioactivity detected in the emulsion. The results obtained by FREMLIN and WALTERS on the double β-decay activity of 16 elements are reported in Table II.

The minimum half-lives were calculated starting from the assumption that an activity equal to one tenth of the total number of disintegrations occurring in a layer of the sample of thickness equal to the β-particle range, would be observed at the surface. Consequentely the calculated values of the

TABLE II.

Sample	Counts in loaded area		Background	Exposure (days)	Activity (β/mm² day)	Min. half-life (years)
	Electrons	α-particles				
$CaCO_3$	86.0± 4.5	none	7.5±1.6	56	0.19±0.07	$7 \cdot 10^{16}$
Cromium	66.0± 4.0	4.3 ± 0.6	66.0±3.0	27	0.00±0.15	$2.2 \cdot 10^{17}$
Iron	76.5± 4.3	1.0 ± 0.4	75.5±1.6	56	0.18±0.07	$1.3 \cdot 10^{17}$
Nickel	52.0± 1.6	0.5 ± 0.1	52.5±1.6	27	0.00±0.07	$3.2 \cdot 10^{17}$
Zinc	44.8± 1.8	0.3 ± 0.1	47.5±1.3	27	0.00±0.12	$2 \cdot 10^{17}$
Germanium	50.0± 1.7	none	52.5±1.6	27	0.00±0.08	$2.8 \cdot 10^{17}$
Molybdenum	87.5± 2.1	none	54.1±1.2	27	1.24±0.09	$1.5 \cdot 10^{16}$
$SrCO_3$	79.1± 3.4	none	63.8±1.5	56	0.27±0.07	$3.4 \cdot 10^{16}$
Cadmium	86.0± 4.6	4.5 ± 1.0	75.5±1.6	56	0.19±0.09	10^{17}
SnO	84.0± 4.8	1.5 ± 0.5	63.8±1.5	56	0.56±0.09	$1.4 \cdot 10^{16}$
^{124}Sn	138 ±7	20.6 ± 1.4	46.4±2.3	15	6.1 ±0.5	$2 \cdot 10^{15}$
Tellurium	77.5± 5.0	none	63.8±1.5	56	0.24±0.11	$1.3 \cdot 10^{16}$
$BaCO_3$	115 ±15	7 ± 1	47.2±2.4	15	4.5 ±0.4	$1.8 \cdot 10^{15}$
Tungsten	53.5± 1.4	0.95± 0.10	39.1±0.9	27	0.54±0.06	$2 \cdot 10^{16}$
Osmium	311 ±20	51 ±10	75.5±1.6	56	40 ±3.6	$2.3 \cdot 10^{14}$
Platinum	67.2± 3.3	3.3 ± 1.0	52.5±1.6	27	0.55±0.13	$1.7 \cdot 10^{16}$

minimum half-lives depend on the values assumed for the total disintegration energy; the values in Table II were calculated for a disintegration energy equal to 2 MeV. For a different energy release the value of the minimum half-life must be multiplied by a factor proportional to the β-particle range corresponding to the assumed energy. If we assume, for instance, a disintegration energy of 4 MeV we should increase the half-lives of Table II by a factor 2.1. The results of FREMLIN and WALTERS are negative, with the exception of the result on molybdenum, regardless of the value assumed for the disintegration energies. The positive result on the molybdenum decay seems to be contradicted by the result of SELIG (pag. 153) on the same transition. No other relevant conclusions can be drawn from the results of Table II because only total β-decay activity was measured without attempt to analyse the possible « spatial coincidence » of the electron tracks.

* * *

The nuclear emulsion technique could be very suitable for searchs for double β-decay, if the sample to be investigated would be incorporated into the emulsion; the « spatial coincidence » of the event could be easily checked in this way. Some attempts of this type have been made by us with rather good results using melted tellurium. Two possibilities were found: the first is to introduce the substance into the emulsion by means of quartz capillary

tubes [50] (glass seems to be less suitable, owing to its background radioactivity); the second method (probably better) consists of evaporating the substance onto small diameter quartz fibers. The wires so obtained can easily be incorporated into the emulsion. With both these methods the « spatial coincidence » of the events can be checked with quite good precision (few microns).

The emulsions can easily be exposed for periods of a few months, without latent-image fading or loss of photographic sensivity. Consequentely very effective shielding is needed in order to bring about a conclusive reduction of the background, because the nuclear emulsion is an « integral » detector. Very interesting investigations were carried out by FREMLIN and WALTERS on the background before [51] and after [52] their experiment on the double β-decay. The background observed in the underground exposure (2.54 ± 0.09 tracks per mm^2 of a 200 μ thick emulsion per day) was ascribed to the thorium and potassium present in glass backing (22% and 39% respectively), to the contamination of the ^{14}C present in the gelatine (33%) and to the owerall contribution due to cosmic radiation and external radioactivity (6%). A very good background reduction was reached by coating the emulsion on a support free from potassium or torium; the background was reduced to only (0.86 ± 0.11) tracks per mm^2 per day in a 200 thick emulsion by using 1.5 mm thick perspex sheets as backing. The authors concluded [52] that the background could be attributed mostly to the ^{14}C present in the gelatine. However, on the basis of a more recent analysis (*) the same authors found that this conclusion could not be completely correct and that some 40% of the background was due to γ-radiation from outside. The sandstone near the mine where FREMLIN and WALTERS carried out their experiments on the background contains a fairly large quantity of thorium; it seems therefore possible to reach a further reduction of the background by a factor of two, by carrying out the experiment in a region free from radioactivity.

At the present time it is quite difficult to obtain good measurements of the energies of the two electrons in double β-decays by using nuclear emulsions; it seems likely, however, that these measurements can be improved by using the new type of high density emulsions. The weakness of the method based on nuclear emulsions lies mainly in the impossibility of checking the « time coincidence » of the events; in many cases a double β-decay can be simulated by a sequence of two single β-decays through the intermediate isotope. This could be the case of the relevant excess radioactivity found by FREMLIN and WALTERS for the molybdenum. The « time coincidence » of the possible double β-decay could be checked by moving the active sample very

(*) Private communication; we are very grateful to Dr. J. H. FREMLIN for having communicated to us their more recent results on the reduction of the background.

slowly inside the emulsion (*); this seems not too difficult if the sample is incorporated in the emulsion in form of thin wires, as suggested before.

3·4. *Experiments with chemical methods.* – Various experiments on the double β-transitions $^{128}_{52}\mathrm{Te} \rightarrow {}^{128}_{54}\mathrm{Xe}$ and $^{130}_{52}\mathrm{Te} \rightarrow {}^{130}_{54}\mathrm{Xe}$ were carried out by M. C. INGHRAM and J. H. REYNOLDS with a method based on the chemical extraction of the daughter element. In their preliminary experiment [53] these authors investigated the amount and the isotopic constitution of xenon accumulated during geological time in a 430 g sample of tellurium ore, containing 12% of tellurobismutite (Bi_2Te_3). The sample, having an estimated age of about $1.5 \cdot 10^9$ years, was crushed and heated in a quartz bottle well above the melting point. The gas evolved was collected and analysed by means of a convectional mass spectrometer calibrated both for chemical and for isotopic analysis with mixtures of argon and xenon. A mixture of $(1.3 \pm 0.3) \cdot 10^{-5}$ cm³ S.T.P. xenon plus $(1.3 \pm 0.3) \cdot 10^{-2}$ cm³ S.T.P. argon was found; the isotopic abundance of xenon is compared in Table III with the

TABLE III. – *Abundance of the xenon isotopes.*

	124	126	128	129	130	131	132	134	136
Xenon in the sample	0.015	0.018	0.25	2.47	0.394	2.00	2.55	1.0	0.847
Natural xenon	0.008 9	0.008 3	0.18	2.49	0.386	2.009	0.558	1.0	0.849

isotopic abundance of natural xenon. All abundance are referred to the abundance of ^{134}Xe placed equal to one.

It will be noted that evidence for radiogenic xenon occours for ^{128}Xe and ^{130}Xe; from this apparently positive result minimum half-lives of $8 \cdot 10^{19}$ years and $1.3 \cdot 10^{19}$ years, respectively, were calculated.

It should be noted, as observed by the authors, that radiogenic xenon may have excaped from the ore since the time of the original mineralization This possibility seems remote however, if one remembers that minerals are able to retain much more diffusible gases, such as helium. It seems more likely that a loss of xenon have occured owing to an hydrothermal alteration of the telluride.

A second experiment [54] was carried out with 371 g of tellurium ore of the same age containing 70% Bi_2Te_3. This sample was obtained from a mine at a depth of 240 m, where hydrothermal alteration of the crystals of telluride was extremely unlikely. A misture of $2.6 \cdot 10^{-7}$ cm³ S.T.P. xenon and $2.4 \cdot 10^{-3}$ cm³ S.T.P. argon was obtained with the same method used in the

(*) Suggestion of Prof. G. P. S. OCCHIALINI.

previous experiment. A considerable excess of the isotopes ^{129}, ^{130} and ^{131}Xe was found in the extracted xenon, for comparison with the isotopic abundance of the natural one. The excess of ^{129}Xe and ^{131}Xe was attributed to (n, γ) reactions on ^{128}Te and ^{130}Te, respectively, from some unknown neutron source. In fact considerable amounts of uranium were found in the immediate neighborhood of the mine from which the tellurobismutite was taken. Part of the excess of ^{129}Xe was also explained as due to single β-decay of ^{129}I (about $3 \cdot 10^7$ years), a nuclide not yet detected in nature. The excess of ^{130}Xe was attributed to a double β-decay of ^{130}Te with a half-life of $1.4 \cdot 10^{21}$ years.

These results were improved by a later experiment of R. J. HAYDEN and M. C. INGHRAM [55] who eliminated two weak points of the earlier experiment, namely the tremendous excess of ^{129}Xe and the uncertainty of the detection efficiency of ^{130}Xe. The xenon separation was repeated on a sample from a mine absolutely free from radioactivity; a ^{128}Xe carrier was used to calculate the detection efficiency of the apparatus for ^{130}Xe. A lower limit of $3.3 \cdot 10^{21}$ years was set for the double β-decay of ^{130}Te. It should be noted however, as pointed out by H. SELIG [31-32], that the excess of ^{130}Xe could perhaps be due to a sequence of two single β-decays $^{130}\text{Te} \rightarrow {}^{130}\text{I} \rightarrow {}^{130}\text{Xe}$. On the other hand an investigation of this transition is very difficult with the present experimental techniques, owing to the extremely low limit of the half-life that can be set by experiments of this type.

A completely different chemical method was employed by C. A. LEVINE *et al.* [56] to detect the possible double β-decay $^{238}_{92}\text{U} \rightarrow {}^{238}_{94}\text{Pu}$. The method consisted in loking in a uranium sample for 90 year ^{238}Pu. 14 kg of 6 year old UO_3 were dissolved in nitritic acid and Pu (IV) was precipitated with lanthanum floride. The plutonium so obtained was oxidated, extracted in diethyl ether and re-extracted into water. After repeating this cycle five times the final sample was analysed with an α-particle detector in order to detect the presence of the 5.51 MeV α-activity of the ^{238}Pu. From the negative result ((0.00 ± 0.1) counts per hour), a minimum half-life of $6 \cdot 10^{18}$ years was obtained, taking into account the chemical and counting efficiency. As pointed out by the authors the results obtained with this chemical method could be strongly improved by using larger sources, containing uranium of geological age, such as pitchblende ore. In this case, however, it must be remembered that a certain amount of ^{239}Pu is contained in the ore.

* * *

From the analysis of the preceeding experiments it can be seen that the chemical methods, although not recent, can set minimum half-lives much longer than the other methods. It is surprising therefore that no recent investigations on the double β-decay have been made with these methods, which

seem so far the only ones suitable to reach results comparable with the theoretical estimates. The chemical methods can detect very low activities owing their tremendous « sensitive time » which is of a factor 10^{10} longer than the measurement time of any other method. On the other side the weakness of these methods lies in the impossibility of detecting the « time and spatial coincidence » of the events. The results obtained with chemical methods can be used as experimental evidence for a value of the half-life only if they are associated with very careful investigations that can rule out the possibility of a sequence of two single β-decay through the intermediate isotope. On a more general basis the possibility that the presence of the daughter isotope might be due to causes other than double β-decay, and, on the other hand, the possibility that part of the daughter isotope might have escaped because of geological alterations of the rocks should be carefully considered.

4. – Mass and nuclear structure of suitable isotopic triplets.

As it is seen from their derivation in Sect. 2 the theoretical estimates for the lifetimes of double β-decays, or equivalent processes, depend rather radically on the energy available in the final states and on the degree of forbidness of the two virtual transitions that « sum up » to give the actual decay. In order to be able, therefore, to compare these theoretical predictions with experimental data, some knowledge has to be gained of the nuclear states from, to and through which the transition takes place.

The information we need is twofold: first, we have to know the energy levels of the initial and final nuclei and also of the accessible states of the intermediate nucleus, since all these quantities appear in the theoretical expression for the transition amplitude. We also need information about the intrinsic angular momentum of those nuclear states which might enter into the process It might be worth reminding here that, for allowed transitions, the change in angular momentum cannot exceed unity (in nucleon spin units). Each degree of forbiddingness, on the other side, gives rise to an increase in the lifetime of the order $10^4 \div 10^5$.

4·1. *Informations about the energy of the nuclear states.* – Let us start with a few remarks on the importance of the study on the energy levels, by means of mass measurements of the suitable isotopic triplets.

First of all, one needs to be sure thet the decay $A \to B + e + \nu$ is energetically forbidden (*); otherwise the process is simply a sequence of two real single

(*) We shall come back briefly later on to the case in which this decay, although energetically allowed, is damped by the large difference in spin between the initial and final nuclear states.

β-decays, and the emission of no neutrinos in the overall process has negligible probability, indipendently of the choice of the theory. Second one should choose, for the experimental detection of the phenomenon, those triplets of isotopes for which the energy release is reasonably high. The reason for this is that the lifetimes decrease rapidly when ε^0 increases, due to the larger volume available in phase space to the leptons which are produced. Finally a detailed knowledge of the mass difference between the initial and final nuclear states is necessary: *a*) to locate the peak in the sum of the energies of the emitted electrons, which has to be expected in a process without the emission of neutrinos: *b*) to determinate the interval within which the sum of the electron energies lies for most of the transitions in which two neutrinos are emitted.

From an analysis of the experiments performed so far it is easily seen how often a poor accuracy in the determination of the masses, and consequentely of the energy of the nuclear states, has led to an incorrect interpretation of the results obtained.

It is well known that the atomic masses can be determined by many different methods, namely by means of mass-spectroscopy, determination of energy balance in nuclear reactions, total β-decay energies, α-disintegration energies and determination of atomic masses by microwave spectroscopy. There is sometimes a discrepancy between the results on mass measurements obtained with different methods; thus data from different sources must be chosen in order to calculate a table of atomic masses. It will be in general reasonable to use the data from nuclear reactions for atoms of low atomic weight ($A < 40$) and to use mass spectroscopy data for heavier atoms.

A recent table of atomic masses was prepared by H. E. DUCKWORTH [59] on the basis of an original work of A. H. WAPSTRA [60]. DUCKWORTH accepted the framework established by mass spectroscopic data; these data were then adjusted in order to make them reasonably compatible with the mass data available from other sources. This table refers to isotopes with $A > 40$; data on isotopes of low atomic weight were obtained from other authors [61-62]. Only ^{36}A of these isotopes can give rise to a double transition.

All the energetically possible double negative β-decays for which the mass of the initial and final nucleus is reasonably well known are reported in Table IV. The percentage of the various isotopes, reported in the second column, was obtained from the review paper of D. STROMINGER *et al.* [63]. The mass differences and the energy releases reported in the third and fourth columns were obtained from DUCKWORT's table; in the last column are reported a few values for energy release calculated on the basis of the results of recent and particulary reliable experiments. All the energy releases were calculated for transitions between the ground states of both the father and the daughter nucleus. The possible transitions for double electron-capture, electron capture-positron emission and positive double β-decay are reported in Table V. The

TABLE IV. – *Possible negative double β-decay.*

Transition	Isotopic aboundance (%)	Mass difference (m MU)	Energy release (MeV)	
			H. E. DUCKWORTH [59]	Others
$^{46}_{20}Ca \rightarrow ^{46}_{22}Ti$	0.0033	0.71 ±0.16	0.661±0.15	0.984±0.004 [62]
$^{48}_{20}Ca \rightarrow ^{48}_{22}Ti$	0.185	4.54 ±0.11	4.227±0.10	4.270±0.006 [62]
$^{70}_{30}Zn \rightarrow ^{70}_{32}Ge$	0.62	1.062±0.150	0.989±0.14	—
$^{76}_{32}Ge \rightarrow ^{76}_{34}Se$	7.67	2.04 ±0.18	1.899±0.17	—
$^{80}_{34}Se \rightarrow ^{80}_{36}Kr$	49.82	0.12 ±0.14	0.112±0.13	0.112±0.019 [64]
$^{82}_{34}Se \rightarrow ^{82}_{36}Kr$	9.19	3.20 ±0.21	2.980±0.195	—
$^{86}_{36}Kr \rightarrow ^{86}_{38}Sr$	17.37	1.44 ±0.21	1.341±0.195	—
$^{94}_{40}Zr \rightarrow ^{94}_{42}Mo$	17.40	1.22 ±0.60	1.136±0.56	—
$^{96}_{40}Zr \rightarrow ^{96}_{42}Mo$	2.80	3.63 ±0.50	3.380±0.465	—
$^{100}_{42}Mo \rightarrow ^{100}_{44}Ru$		9.62		2.3 ±0.2 [58]
$^{104}_{44}Ru \rightarrow ^{104}_{46}Pd$	18.5	1.21 ±0.32	1.127±0.298	–
$^{110}_{46}Pd \rightarrow ^{110}_{48}Cd$	12.65±0.85	1.59 ±0.57	1.480±0.53	—
$^{114}_{48}Cd \rightarrow ^{114}_{50}Sn$	28.86	0.61 ±0.50	0.568±0.465	0.559±0.028 [64]
$^{116}_{48}Cd \rightarrow ^{116}_{50}Sn$	7.58	2.80 ±0.46	2.607±0.430	—
$^{122}_{50}Sn \rightarrow ^{122}_{52}Te$	4.71	0.42 ±0.43	0.390±0.401	0.42 [63]
$^{124}_{50}Sn \rightarrow ^{124}_{52}Te$	5.98	2.48 ±0.42	2.31 ±0.39	—
$^{128}_{52}Te \rightarrow ^{128}_{54}Xe$	31.79	0.88 ±0.71	0.819±0.662	—
$^{130}_{52}Te \rightarrow ^{130}_{54}Xe$	34.49	3.15 ±0.40	2.933±0.37	—
$^{134}_{54}Xe \rightarrow ^{134}_{56}Ba$	10.44	1.01 ±0.14	0.94 ±0.13	1.080±0.088 [65]
$^{136}_{54}Xe \rightarrow ^{136}_{56}Ba$	8.87	2.870±0.093	2.672±0.087	2.643±0.087 [65]
$^{142}_{58}Ce \rightarrow ^{142}_{60}Nd$	11.07	1.71 ±0.12	1.590±0.112	1.685±0.071 [65]
$^{148}_{60}Nd \rightarrow ^{148}_{62}Sm$	5.71±0.051	2.04 ±0.21	1.899±0.195	
$^{150}_{60}Nd \rightarrow ^{150}_{62}Sm$	5.60±0.05	3.92 ±0.10	3.650±0.093	
$^{154}_{62}Sm \rightarrow ^{154}_{64}Gd$	22.61±0.367	0.97 ±0.38	0.903±0.354	
$^{160}_{64}Gd \rightarrow ^{160}_{66}Dy$	21.75±0.15	1.50 ±1.00	1.396±0.931	
$^{238}_{92}U \rightarrow ^{238}_{94}Pu$	99.275			1.0 [63]

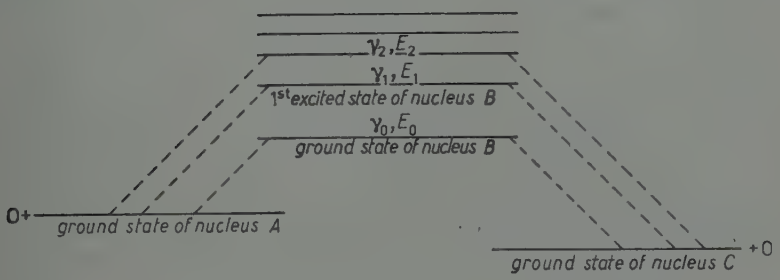
TABLE V. – *Possible double electron captures, electron captures-positron emissions and positive double β-decays.*

Transition	Isotopic abundance (%)	Mass difference (mMU)	Energy release (MeV)	
			H. E. DUCKWORTH	Others
$^{36}_{18}A \rightarrow {}^{36}_{16}S$	0.337			0.486 [63]
$^{40}_{20}Ca \rightarrow {}^{40}_{18}A$	96.97	0.242±0.030	0.225±0.028	0.170±0.010 [63]
$^{50}_{24}Cr \rightarrow {}^{50}_{22}Ti$	4.31	1.15 ±0.11	1.071±0.102	1.175±0.003 [62]
$^{54}_{26}Fe \rightarrow {}^{54}_{24}Cr$	5.84	0.680±0.076	0.633±0.071	
$^{58}_{28}Ni \rightarrow {}^{58}_{26}Fe$	67.76	2.036±0.012	1.896±0.011	1.896±0.007 [64]
$^{64}_{30}Zn \rightarrow {}^{64}_{28}Ni$	48.89	1.187±0.007	1.105±0.006	1.053±0.006 [64]
$^{74}_{34}Se \rightarrow {}^{74}_{36}Ge$	0.87	1.31 ±0.18	1.22 ±0.17	1.192±0.019 [64]
$^{78}_{36}Kr \rightarrow {}^{78}_{34}Se$	0.354	2.81 ±0.21	2.616±0.195 (0.572±0.195)	
$^{84}_{38}Sr \rightarrow {}^{84}_{36}Kr$	0.555±0.005	1.74 ±0.21	1.620±0.195	1.74 [63]
$^{92}_{42}Mo \rightarrow {}^{92}_{40}Zr$	15.86	1.72 ±0.57	1.60 ±0.53	
$^{96}_{44}Ru \rightarrow {}^{96}_{42}Mo$	5.59±0.048	3.02 ±0.50	2.812±0.465 0.768±0.465	
$^{102}_{46}Pd \rightarrow {}^{102}_{44}Ru$	0.88±0.08	1.30 ±0.32	1.210±0.298	1.11 [63]
$^{106}_{48}Cd \rightarrow {}^{106}_{46}Pd$	1.22	2.86 ±0.50	2.663±0.465 (0.619±0.465)	
$^{112}_{50}Sn \rightarrow {}^{112}_{48}Cd$	0.95	2.02 ±0.50	1.881±0.465	1.862±0.009 [64]
$^{120}_{52}Te \rightarrow {}^{120}_{50}Sn$	0.089	1.97 ±0.42	1.834±0.39	
$^{124}_{54}Xe \rightarrow {}^{124}_{52}Te$	0.096	2.92 ±0.42	2.719±0.39 (0.675±0.39)	
$^{126}_{54}Xe \rightarrow {}^{126}_{52}Te$	0.090	0.96 ±0.42	0.894±0.39	0.968±0.019 [64]
$^{130}_{56}Ba \rightarrow {}^{130}_{54}Xe$	0.115±0.015	2.74 ±0.05	2.551±0.046 (0.507±0.046)	2.533±0.093 [64] (0.509±0.046)
$^{132}_{56}Ba \rightarrow {}^{132}_{54}Xe$	0.144±0.046	0.89 ±0.13	0.829±0.121	
$^{136}_{58}Ce \rightarrow {}^{136}_{56}Ba$	0.193	2.73 ±0.22	2.542±0.205 (0.498±0.205)	
$^{138}_{58}Ce \rightarrow {}^{138}_{56}Ba$	0.250	1.140±0.215	1.061±0.200	1.19 [63]
$^{144}_{62}Sm \rightarrow {}^{144}_{60}Nd$	3.11±0.037	1.95 ±0.14	1.815±0.130	1.722±0.130 [65]

energy releases in the last two columns were calculated for double electron-capture; the energy releases for positive double β-decay are reported in brakets when this transition is energetically possible.

4·2. *Nuclear structure.* – It is obvious that the determination of the energy of the nuclear states involved in a particular double β-decay is generally not sufficient to estimate the transition amplitude and therefore the lifetime to be expected according to the different theories. A very important role might be played by the degree of forbiddingness of the transition; it is therefore essential a knowledge of the intrinsic angular momentum, I, of the nuclear levels from, to and through which the decay takes place.

Since the transition probability for double β-decay is exceedingly small even in the most favorable conditions, it will be wise, for example, to perform experiments only when the three isotopes in consideration are such that the spin of the initial nuclear state, I_A, differs by a factor of at most two from the spin I_C of one of the energetically accessible levels of the nucleus C. In fact, if $I_A - I_C > 2$, at least one of the intermediate transitions is forbidden, and the expected lifetime becomes exceedingly long. In what follows we shall restrict ourselves to those cases in which the inequality $I_A - I_C \leqslant 2$ is satisfied. In fact we shall study only double β-decays in which the transition takes place between two spin zero states (*). This situation is typical of the most common transitions in which both A and C are even-even nuclei; their ground states have then spin zero and positive parity. Since the decay takes place into the ground state of nucleons C, the transition is then particulary favored also energetically. The situation we are considering is visualized by the following scheme:



In the scheme $\gamma_0, \gamma_1, \gamma_2, \ldots$ stand for the spin (and parity) assignements, $E_0, E_1, E_2, \ldots$ for the energies of the corresponding states. A transition to

(*) All what we are going to say can be adapted in an obvious way to the general case $0 \leqslant I_A - I_C \leqslant 2$.

some excited state of nucleus C may be allowed in certain cases; for the sake of simplicity, and without essential loss of generality, we neglect this possibility (*).

Neglecting transitions forbidden of order $\geqslant 1$, the double β-decay A → C goes, virtually, through those states of the nucleus B which have spin $\leqslant 1$. The ground state of nucleus B shall not, and in most of the actual cases does not, belong to these states. A simple remark presents itself at this point. The rough estimate of the matrix elements for the process was obtained (see Sect. 2) on the ground of the experimental data for the corresponding single β transitions, in our case, for the two decays $B_{\text{ground state}} \to A_{\text{ground state}}$ and $B_{\text{ground state}} \to C_{\text{ground state}}$. No such experimental source is available for the matrix elements $B_{n\text{-th excited state}} \to A_{\text{ground state}}$. Lacking anything better, it is assumed here that the estimates given in Sect. 2 are still valid; this is, after all, not worst of the assumptions made so far.

Let us assume that the n-th excited state is the first, on the energy scale, which can be reached from $A_{\text{ground state}}$ through an allowed transition. Its excitation energy, $E_n - E_0$, might be of importance in estimating the transition amplitude. It is at this point useful to treat separately the two cases of double β-decay with and without neutrinos.

a) Decay without neutrinos. As we have seen in Sect. 2, the most important contribution is most likely to come from intermediate states which contain a virtual neutrino with energy ~ 30 MeV. The approximation introduced was then to neglect the difference in energy between the initial and intermediate nuclear states in the energy denominator (see *e.g.* eq. 2.8). As far as the excitation energy, $E_n - E_0$, of the n-th state of nucleus B is small compared to 30 MeV (and this seems to be the case in all actual situations), the approximation keeps its validity and the predicted values for the lifetime do not change. We conclude therefore that the estimated lifetime for a neutrinoless double β-decay between two even-even nuclei is somewhat insensitive to the spin assignment of the ground state of the intermediate nucleus, at least in so far as the excitation energy of the first level of nucleus B which has spin 1 or 0, is smaller than, say, 5 MeV.

b) Decay with two neutrinos. Since the maximum energy available to electrons and neutrinos depends only on the relative energies of $A_{\text{ground state}}$ and $C_{\text{ground state}}$ and is independent of the « intermediate » nuclear state, the phase space available does not change. We already agreed not to change

(*) Strictly speaking, one should also take into account the possibility of a decay from an excited state of nucleus A; this contribution is, however, non-negligible only when the β-decay $A_{\text{exc.}} \to B + e + \nu$ competes with a highly forbidden γ-transition from the isomeric state $A_{\text{exc.}}$ to A_{ground}.

the estimates of the nuclear matrix elements; the only part of the transition amplitude which is sensible to the energy of the n-th level of nucleus N is therefore the energy denominator. The approximation was made in Sect. 2 to replace it with a constant. This constant is given, roughly speaking, by one-half of the energy release, plus one electron mass, plus the average energy distance between the ground state of nucleus A and those levels of nucleus B which contribute appreciably to the reaction. The actual value chosen for the constant depends of course on the excitation energy of the state $B_{n\text{-th}}$, and the more so the smaller the term $m_e + E_{B_{\text{ground state}}} - \frac{1}{2}(E_{A_{\text{ground state}}} + E_{C_{\text{ground state}}})$ is.

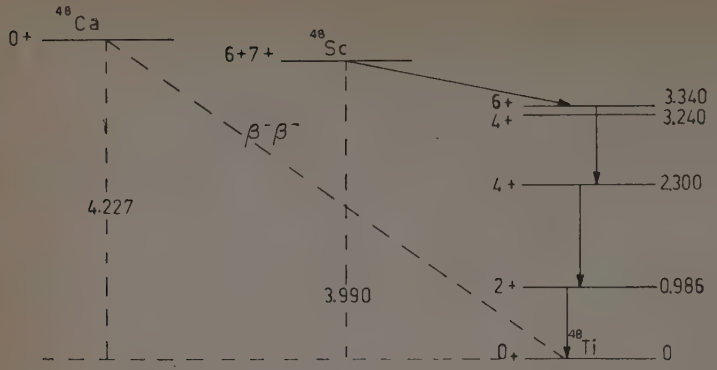
We also recall that, in the final expression for the lifetime, the factor $(E_B - E_{B_{\text{ground state}}}) + m_e + E_{B_{\text{ground state}}} - \frac{1}{2}(E_{A_{\text{ground state}}} + E_{C_{\text{ground state}}})$ appears square so that if, *e.g.*, it is increased by a factor ~ 3 the expected lifetimes become longer by a factor ~ 10. In a decay with two neutrinos, therefore, the lifetime depends sensibly on the spin assignment and the energy spacing of the states of the intermediate nucleus.

4·3. *Possible double* β*-transitions.* – The present-day knowledge of the nuclear states for the isotopic triplets more suitable for negative double β-decay is summarized in Fig. 12. In Fig. 13 nuclear schemes are given for possible double electron capture (2EC), electron capture-positron emission (β⁺EC) and positive double β-decay. For sake of simplicity only the more suitable isotopic triplets were studied (energy release > 1 MeV); for each triplet are reported the references from which the nuclear data were obtained.

It might be instructive to consider in some more detail a few double β transitions particularly interesting from the theoretical point of wiew and experimentally promising.

1) ${}^{48}_{20}\text{Ca} \to {}^{48}_{21}\text{Sc} \to {}^{48}_{22}\text{Ti}$ (similar considerations can be drawn for the triplet 12 *f*)).

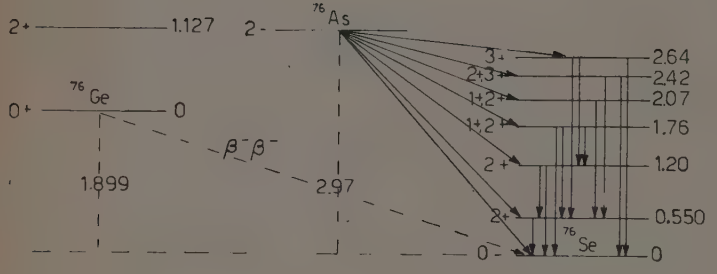
The decay of ⁴⁸Ca into ⁴⁸Sc is energetically allowed; this transition has however never been detected [31-32, 66]. Experimental evidence is for a lifetime smaller than $2 \cdot 10^{16}$ years. This implies that the transition is highly forbidden, a result also obtainable from the remark that the decay ${}^{48}_{21}\text{Sc} \to {}^{48}_{22}\text{Ti}$ takes place almost entirely to a state of ⁴⁸Ti which has spin 6 and parity + and not to a lower level characterized by spin 4 and parity +. This is consistent with an assignment 6+ or 7+ to the ground state of ⁴⁸Sc and in turn together with the assignment 0+ to the ground state of ⁴⁸Ca, implies that the decay ${}^{48}_{20}\text{Ca} \to {}^{48}_{21}\text{Sc}$ is forbidden of order 5 or 6. Using the nuclear matrix elements one may tentatively predict a value of order 10^{20} years fo the single β-decay of ⁴⁸Ca. If this estimate were correct, the triplet of isotopes we are discussing might be rather convenient for the study of double β-decay, because



$^{48}_{20}Ca \rightarrow {}^{48}_{21}Sc \rightarrow {}^{48}_{22}Ti$.

Fig. 12*a*.

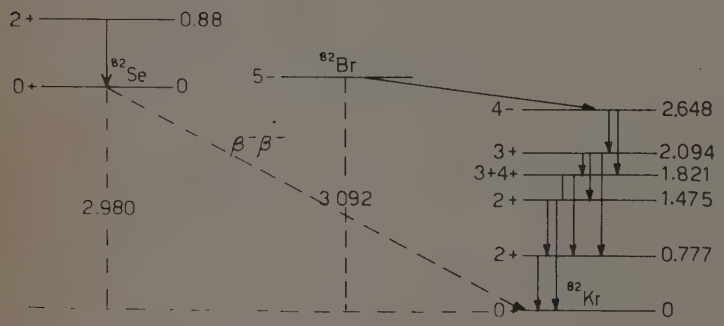
Ref. [63, 67, 68].



$^{76}_{32}Ge \rightarrow {}^{76}_{33}As \rightarrow {}^{76}_{34}Se$.

Fig. 12*b*.

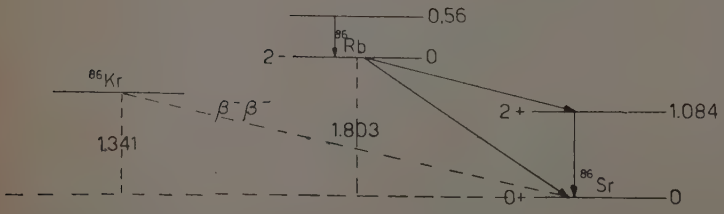
Ref. [63, 69, 70]



$^{82}_{34}Se \rightarrow {}^{82}_{35}Br \rightarrow {}^{82}_{36}Kr$.

Fig. 12*c*.

Ref. [63, 71, 73].

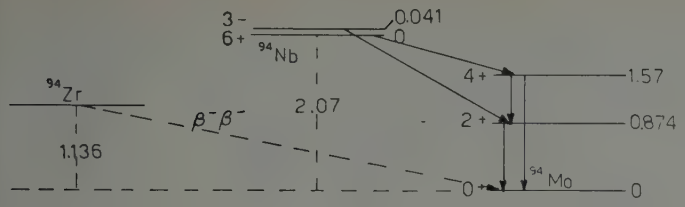


$^{86}_{36}Kr \rightarrow {}^{86}_{37}Rb \rightarrow {}^{8}_{38}Sr$.

Fig. 12*d*.

Ref. [63, 74, 75].

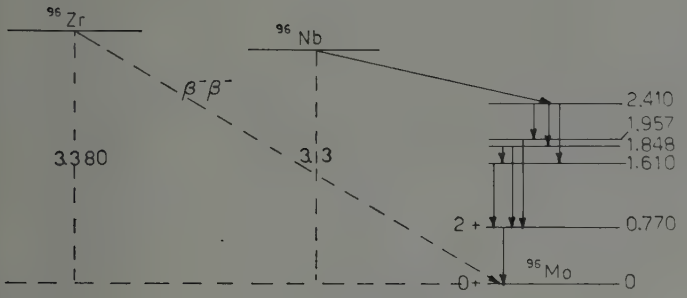
Fig. 12. – Nuclear schemes of the more suitable triplets for negative double β-decay.



$^{94}_{90}Zr \rightarrow {}^{94}_{91}Nb \rightarrow {}^{94}_{92}Mo$

Fig. 12*e*.

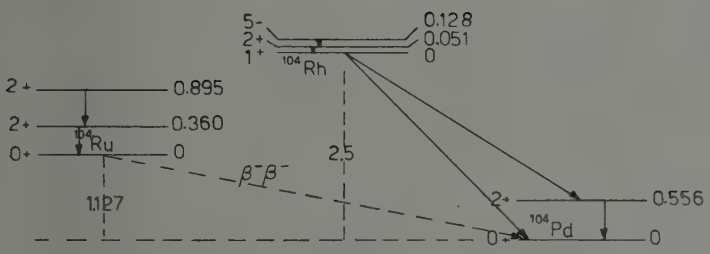
Ref. [63, 76].



$^{96}_{40}Zr \rightarrow {}^{96}_{41}Nb \rightarrow {}^{96}_{42}Mo$.

Fig. 12*f*.

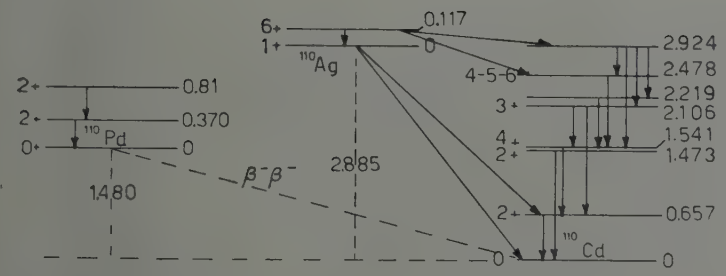
Ref. [63].



$^{104}_{44}Ru \rightarrow {}^{104}_{45}Rh \rightarrow {}^{104}_{46}Pd$.

Fig. 12*g*.

Ref. [63].

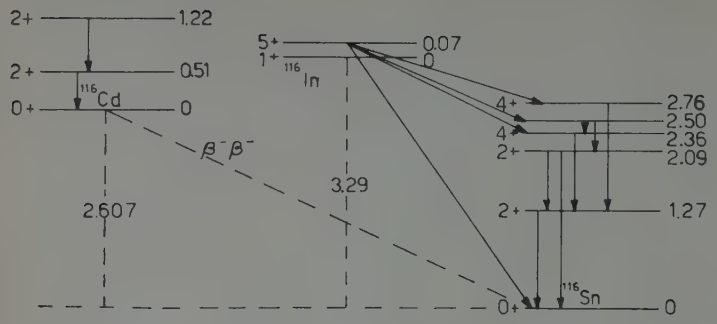


$^{110}_{46}Pd \rightarrow {}^{110}_{47}Ag \rightarrow {}^{110}_{48}Cd$.

Fig. 12*h*.

Ref. [63, 77, 78].

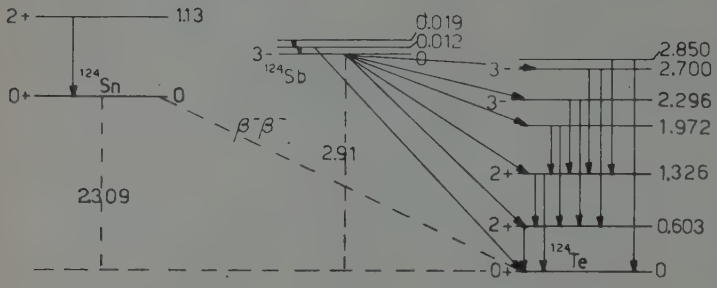
Fig. 12. – (*continued*).



$^{116}_{48}Cd \rightarrow {}^{116}_{49}In \rightarrow {}^{116}_{50}Sn$.

Fig. 12*i*.

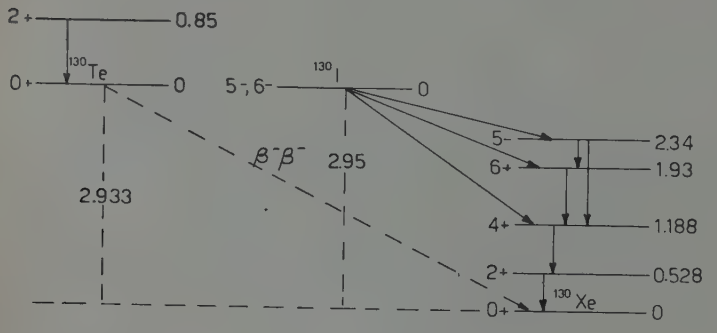
Ref. [63, 73].



$^{124}_{50}Sn \rightarrow {}^{124}_{51}Sb \rightarrow {}^{124}_{52}Te$.

Fig. 12*l*.

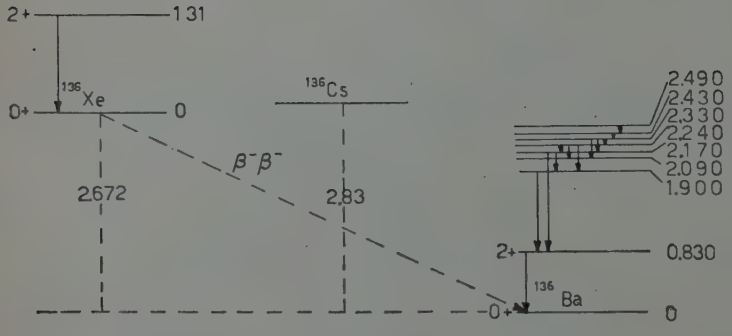
Ref. [63, 79].



$^{130}_{52}Te \rightarrow {}^{130}_{53}I \rightarrow {}^{130}_{54}Xe$.

Fig. 12*m*.

Ref. [63, 80-82].

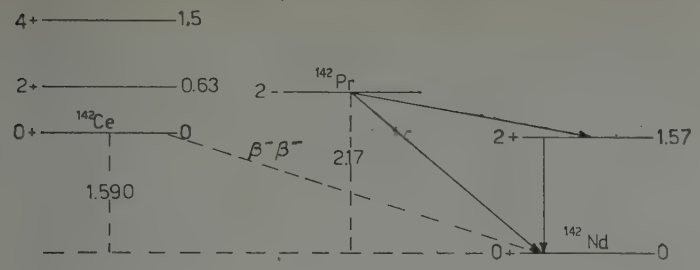


$^{136}_{54}Xe \rightarrow {}^{136}_{55}Cs \rightarrow {}^{136}_{56}Ba$.

Fig. 12*n*.

Ref. [63, 83, 84].

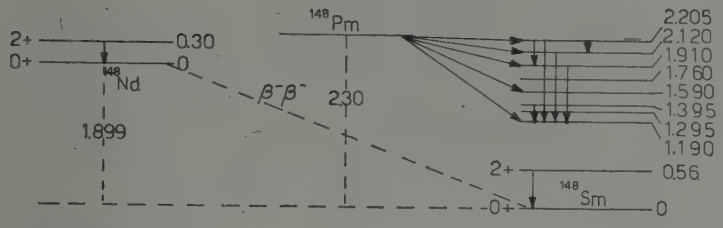
Fig. 12. – (*continued*).



$^{142}_{58}Ce \rightarrow {}^{142}_{59}Pr \rightarrow {}^{142}_{60}Nd$.

Fig. 12*o*.

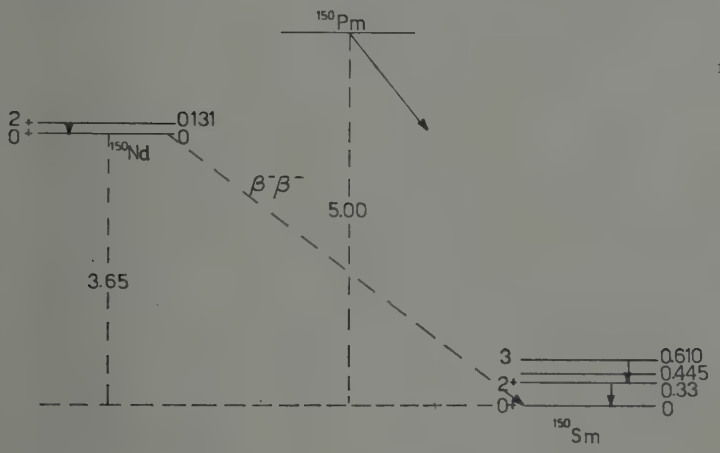
Ref. [63, 85, 86].



$^{148}_{60}Nd \rightarrow {}^{148}_{61}Pm \rightarrow {}^{148}_{62}Sm$.

Fig. 12*p*.

Ref. [63, 87].



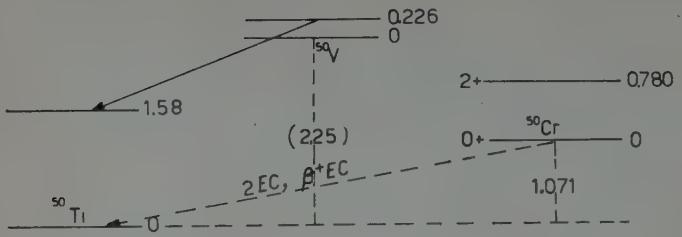
$^{150}_{60}Nd \rightarrow {}^{150}_{61}Pm \rightarrow {}^{150}_{62}Sm$.

Fig. 12*q*.

Ref. [63, 88].

Fig. 12. – (*continued*).

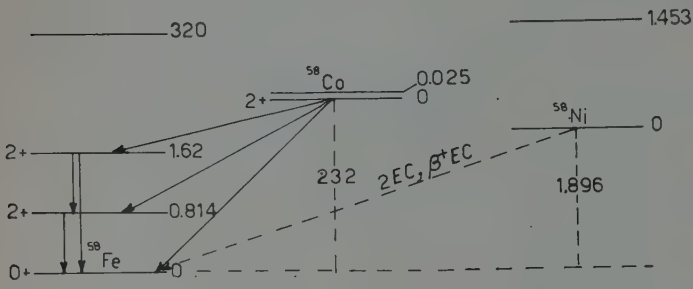
the energy release (4.27 MeV) is considerably large and the states of ^{48}Sc through which the virtual transitions take place have an energy presumably not much higher (say, 2 or 3 MeV) than the energy of the ground state of ^{48}Ca. From the formula given in Sect. 2 it is easily seen that the theoretical estimates for the lifetime of this double β-decay are: $\tau_{no} \sim 2.6 \cdot 10^{15 \pm 2}$ years for no-neutrino emission and with the choice of $\alpha = 1$, which will be made also in the following, and $\tau_{two} \sim 8 \cdot 10^{19 \pm 2}$ years for emission of two neutrinos. It should however be clear that a result $\tau \sim 10^{16}$ years for the lifetime would not represent by itself (*i.e.* without coincidence or correlation measurements) an in-



$^{50}_{24}Cr \rightarrow ^{50}_{23}V \rightarrow ^{50}_{22}Ti$.

Fig. 13*a*.

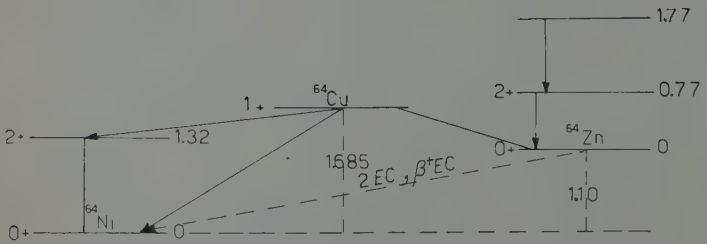
Ref. [89, 90].



$^{58}_{28}Ni \rightarrow ^{58}_{27}Co \rightarrow ^{58}_{26}Fe$.

Fig 13. *b*.

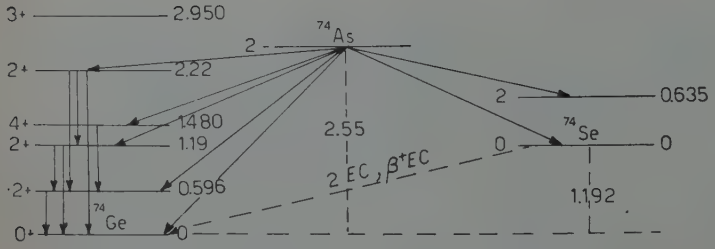
Ref. [63, 91-93].



$^{64}_{30}Zn \rightarrow ^{64}_{29}Cu \rightarrow ^{64}_{28}Ni$.

Fig. 13*c*.

Ref. [63, 94-96].

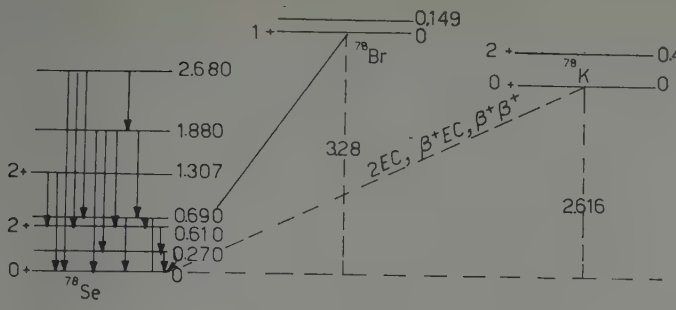


$^{74}_{34}Se \rightarrow ^{74}_{33}As \rightarrow ^{74}_{32}Ge$.

Fig. 13*d*.

Ref. [63, 97-99].

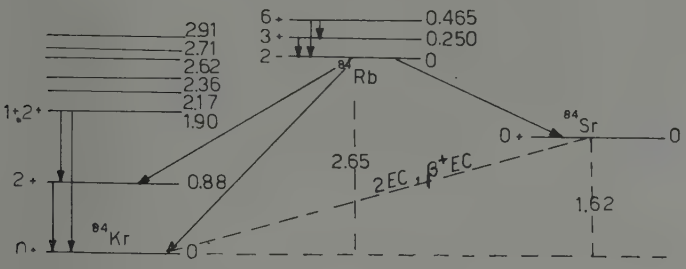
Fig. 13. – Nuclear schemes for positive double β-decay ($\beta^+\beta^+$); double electron capture (2EC) and electron capture, positron emission (β^+EC).



$^{78}_{36}\mathrm{Kr} \rightarrow {}^{78}_{35}\mathrm{Br} \rightarrow {}^{78}_{34}\mathrm{Se}$.

Fig. 13*e*.

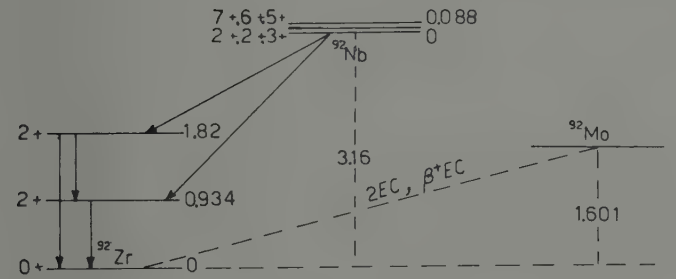
Ref. [63, 72, 100, 101].



$^{84}_{38}\mathrm{Sr} \rightarrow {}^{84}_{37}\mathrm{Rb} \rightarrow {}^{84}_{35}\mathrm{Kr}$.

Fig. 13*f*.

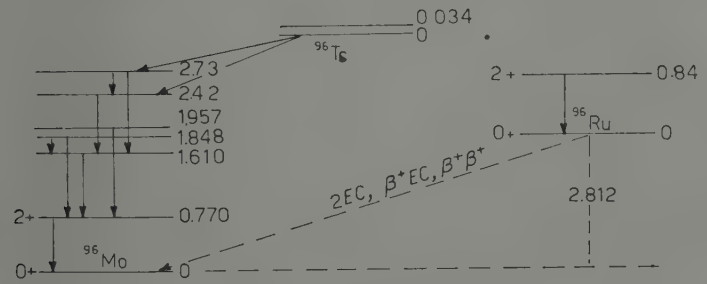
Ref. [63, 72, 102-104].



$^{92}_{40}\mathrm{Mo} \rightarrow {}^{92}_{41}\mathrm{Nb} \rightarrow {}^{92}_{39}\mathrm{Zr}$.

Fig. 13*g*.

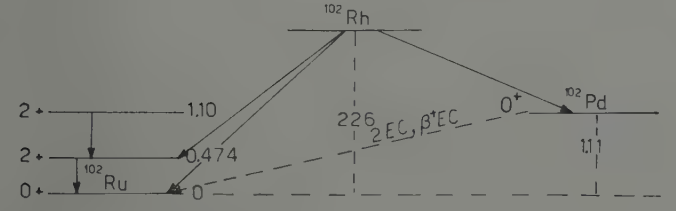
Ref. [63, 100, 105.]



$^{96}_{44}\mathrm{Ru} \rightarrow {}^{96}_{43}\mathrm{Tc} \rightarrow {}^{96}_{42}\mathrm{Mo}$.

Fig. 13*h*.

Ref. [63].

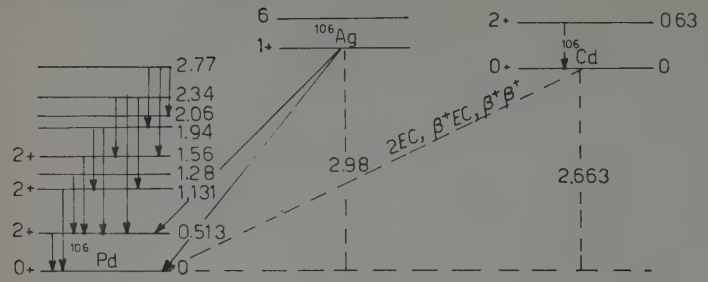


$^{102}_{46}\mathrm{Pd} \rightarrow {}^{102}_{45}\mathrm{Rh} \rightarrow {}^{102}_{44}\mathrm{Ru}$.

Fig. 13*i*.

Ref. [63].

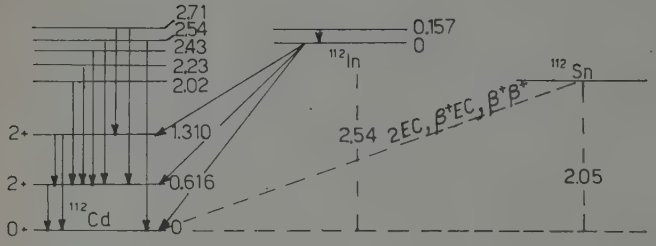
Fig. 13. – (*Continued*).



$^{106}_{48}\mathrm{Cd} \rightarrow {}^{106}_{47}\mathrm{Ag} \rightarrow {}^{106}_{46}\mathrm{Pd}$.

Fig. 13*l*.

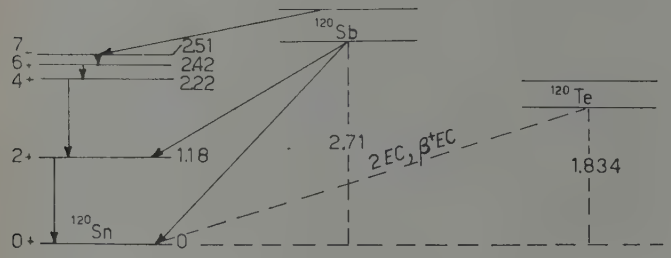
Ref. [63, 106-109].



$^{112}_{50}\mathrm{Sn} \rightarrow {}^{112}_{49}\mathrm{In} \rightarrow {}^{112}_{48}\mathrm{Cd}$.

Fig. 13*m*.

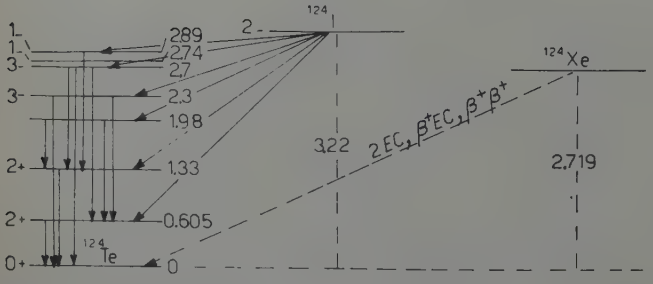
Ref. [63, 110-112].



$^{120}_{52}\mathrm{Te} \rightarrow {}^{12\ 0}_{5\ 1}\mathrm{Sb} \rightarrow {}^{120}_{50}\mathrm{Sn}$.

Fig. 13*n*.

Ref. [63 113-116].

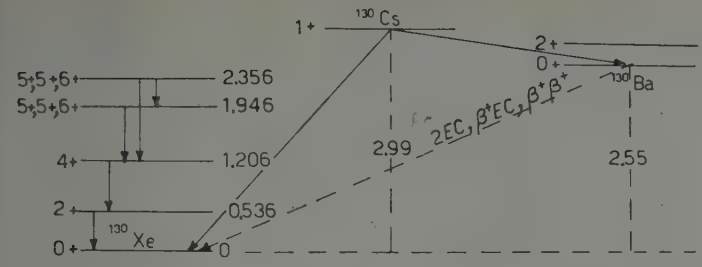


$^{124}_{54}\mathrm{Xe} \rightarrow {}^{124}_{53}\mathrm{I} \rightarrow {}^{124}_{52}\mathrm{Te}$.

Fig. 13*o*.

Ref. [63, 79, 117].

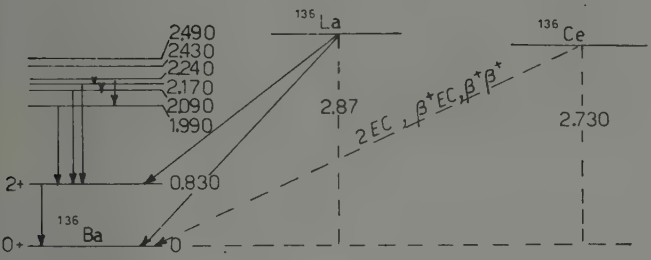
Fig. 13. – (*Continued*).



$^{130}_{56}Ba \rightarrow {}^{130}_{55}Cs \rightarrow {}^{130}_{54}Xe$.

Fig. 13*p*.

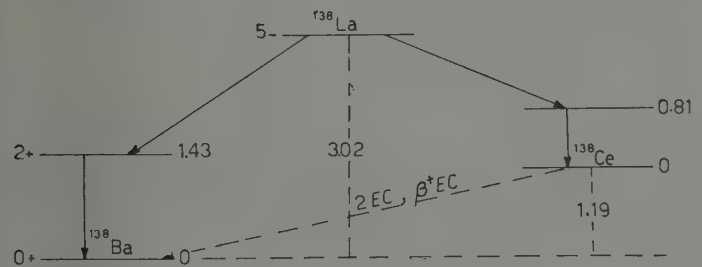
Ref. [63, 82].



$^{136}_{58}Ce \rightarrow {}^{136}_{57}La \rightarrow {}^{136}_{56}Ba$.

Fig. 13*q*.

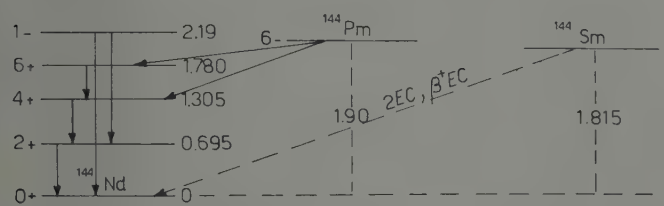
Ref. [63, 83].



$^{138}_{58}Ce \rightarrow {}^{138}_{57}La \rightarrow {}^{138}_{56}Ba$.

Fig. 13*r*.

Ref. [63].



$^{144}_{62}Sm \rightarrow {}^{144}_{61}Pm \rightarrow {}^{144}_{60}Nd$.

Fig. 13*s*.

Ref. [63, 118-120].

Fig. 13. – (*Continued*).

dication in favor of a decay without neutrinos, since the estimate previously given of the lifetime for the decay $^{48}_{20}Ca \rightarrow {}^{48}_{21}Sc$ might be grossly wrong.

2) $^{76}_{32}Ge \rightarrow {}^{76}_{33}As \rightarrow {}^{76}_{34}Se$ (similar consideration can be drawn for triplets 12 *e*) *g*) *h*) *i*) *o*) *r*) and 13 *a*) *d*) *e*) *f*) *g*) *i*) *m*) *n*) *q*) *r*)).

In this case the first intermediate process is certainly virtual (*i.e.* energetically forbidden). From the fairly well known decay scheme of $^{76}_{33}As \rightarrow {}^{76}_{31}Se$

one can conclude that the ground state of ^{76}As is 2 — and that the energy difference between the ground state of ^{76}As and ^{76}Se is 2.97 MeV. Spectroscopic determinations give, for the difference in energy between the ground states of ^{76}Ge and ^{76}Se, a value of $(1.9 \pm .2)$ MeV. The excited states of ^{76}As through which the transitions go, have presumably not very high excitation energy; its average will be denoted by ε. The estimates for the lifetimes are, in a process without emission of neutrinos, $\tau_{no} \sim 1 \cdot 10^{17 \pm 2}$ years and, when two neutrinos are emitted, $\tau_{two} \sim 3 \cdot 10^{21 \pm 2} \cdot (\varepsilon_1 + 5)^2$ years, where ε_1 is given in units in which $m_e = \hbar = c = 1$. To get an idea of the dependence of τ_{two} on the excitation energy of the accessible intermediate levels (we have already seen that τ_{no} is pratically independent of this quantity, at least as for as $\varepsilon_1 \ll 30$ MeV) we choose, as typical, the values $\varepsilon_1 = 0$ and $\varepsilon_1 = 5$ (units $m_e = \hbar = c = 1$). The corresponding estimates for the lifetimes are then

$$\tau_{two,\varepsilon_1=0} \sim 7.5 \cdot 10^{22 \pm 2}; \qquad \tau_{two,\varepsilon_1=5} \sim 3 \cdot 10^{23 \pm 2} \text{ years}.$$

3) Let us finally comment briefly on two processes which might be of interest because of the comparatively short estimated lifetimes:

(*a*) $$^{82}_{34}\text{Se} \to {}^{82}_{35}\text{Br} \to {}^{82}_{36}\text{Kr}.$$

The energy difference between the ground states of ^{82}Br and ^{82}Se is determined to be (0.11 ± 0.19) MeV; this means that one cannot exclude that the decay $^{82}_{34}\text{Se} \to {}^{82}_{35}\text{Br}$ be energetically allowed. The ground state of ^{82}Br is however known, from the scheme of its decay into $^{82}_{36}$Kr, to have spin 5 and parity minus. The transition $^{82}_{34}\text{Se}_{\text{ground state}} \to {}^{82}_{35}\text{Br}_{\text{ground state}}$ is therefore highly forbidden; the further damping due to the smallness of the phase space available makes the probability of this decay negligible. The estimated lifetimes for the double β-decay are

$$\tau_{no} \sim 1 \cdot 10^{16 \pm 2} \qquad \tau_{two} \sim 6 \cdot 10^{21 \pm 2},$$

(we have assumed $\varepsilon_1 = 3$, the result not being very sensitive to this choice).

(*b*) $$^{116}_{48}\text{Cd} \to {}^{116}_{49}\text{In} \to {}^{116}_{50}\text{Sn}.$$

The decay $^{116}_{48}\text{Cd} \to {}^{116}_{49}\text{In}$ is energetically forbidden. The ground state of ^{116}In has presumibly spin 1 and parity plus. This process provides us, therefore, with an example of a virtual transition going through the ground state of the intermediate nucleus. The energy release for the double β-decay is 2.607 MeV; using this value the theoretical estimates for the lifetimes are $1.4 \cdot 10^{16 \pm 2}$ years and $3 \cdot 10^{21 \pm 2}$ years.

5. – Conclusions.

A comparison of the experimental results with the theoretical predictions for the double β-transitions so far investigated is reported in Table VI and Table VII for negative and positive double β-transitions respectively. The theoretical values for the half-lives were calculated from equations (2.12) and (2.16). From this comparison it can be seen that no conclusive experimental evidence can be claimed in favour, or against a double β-decay without emission of neutrinos. Only the experiment carried out by R. J. HAYDEN and M. C. INGHRAM gives a minimum half-life that can be compared with the theory; this result, which was only weakly confirmed by the other negative experiments seems not compatible with the maximal non-conservation of neutrinos. Unfortunately the result is not recent and should be checked with a more accurate one; as we have seen before, the weakness of this experiment lies in the impossibility of measuring the sum of electron energies and of ruling out the alternative possibility of the sequence of two single β-decays.

The difficulties of the comparison between theoretical predictions and experimental results can be summarized in this way:

a) Theoretically there is a fairly large uncertainty (by a factor $10^{\pm 2}$ or more) in the estimates for the lifetimes. This uncertainty comes from the evaluation of nuclear matrix elements; it is therefore indipendent of the choice of the numerical value of the parameters which appear in the Hamiltonian density. Moreover, as we have already noticed, the two extreme choices corresponding to conservation and maximal non conservation of neutrinos, respectively, give only an upper, or, respectively, lower bound to the lifetimes to be expected theoretically. Any intermediate value can be predicted by a proper (but not unique) choice of the parameters. Due to the above-mentioned large uncertainty in the theoretical estimates, experimental data on double β-decay can therefore possibly rule out some choices (*e.g.* a Majorana theory), but cannot give, by themselves, indications as to which of the remaining is the correct one. In this respect, namely to determine the phenomenological Hamiltonian density that should be used to describe the weak interactions of nucleons, electrons and neutrons, double β-decay seems a less suited phenomenon than, say, angular correlations or electron polarization.

b) From the experimental point of view it seems quite difficult to obtain a reliable and conclusive improvement in the measurement of the value of the half-life. The limitation on the detectability of very weak activities with the various experimental techniques was investigated in Sect. **3**; it seems reasonable to draw the conclusion that a much more conclusive result on the lower limit for the half-life can be obtained only with new experimental pro-

TABLE VI. – *Comparison between experimental and theoretical results on the negative double β-decay.*

Transition	Authors	Experimental method	Experimental result	Experimental half-life	Theor. half-life (years) without neutrinos	Theor. half-life (years) with neutrinos
$^{48}_{20}Ca \rightarrow ^{48}_{22}Ti$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 7 \cdot 10^{16}$	$2.6 \cdot 10^{15\pm2}$	$8 \cdot 10^{19\pm2}$
	MC CARTHY	Scintillation counters	Positive?	$1.6 \cdot 10^{17}$		
	AWSCHALOM	Scintillation counters	Negative	$> 2 \cdot 10^{18}$		
	DOBROKHOTOV *et al.*	Scintillation counters	Negative	$> 7 \cdot 10^{18}$		
$^{76}_{32}Ge \rightarrow ^{76}_{34}Se$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 2.8 \cdot 10^{17}$	$1 \cdot 10^{17\pm2}$	$8.5 \cdot 10^{22\pm2}$
$^{96}_{40}Zr \rightarrow ^{96}_{42}Mo$	MC CARTHY	Scintillation counters	Positive?	$6 \cdot 10^{16}$	$4.6 \cdot 10^{15\pm2}$	$2 \cdot 10^{20\pm2}$
	AWSCHALOM	Scintillation counters	Negative	$> 5 \cdot 10^{17}$		
$^{100}_{42}Mo \rightarrow ^{100}_{44}Ru$	FREMLIN and WALTERS	Nuclear emulsion	Positive?	$1.5 \cdot 10^{16}$	$3 \cdot 10^{16\pm2}$	$6 \cdot 10^{21\pm2}$
	WINTER	Cloud chamber	Negative	$> 3 \cdot 10^{17}$		
$^{110}_{46}Pd \rightarrow ^{110}_{48}Cd$	WINTER	Cloud chamber	Negative	$> 1.1 \cdot 10^{18}$	$2.6 \cdot 10^{17\pm2}$	$1 \cdot 10^{24\pm2}$
$^{116}_{48}Cd \rightarrow ^{116}_{50}Sn$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 10^{17}$	$1.4 \cdot 10^{16\pm2}$	$3 \cdot 10^{21\pm2}$
	DATOEUF *et al.*	Scintillation counters	Negative	$> 7 \cdot 10^{16}$		
	WINTER	Cloud chamber	Negative	$> 1 \cdot 10^{17}$		
$^{124}_{50}Sn \rightarrow ^{124}_{52}Te$	FIREMAN	G.M. counters	Positive?	$6.5 \cdot 10^{15}$	$2.5 \cdot 10^{16\pm2}$	$7 \cdot 10^{21\pm2}$
	LAWSON	Cloud chamber	Negative	$> 5 \cdot 10^{16}$		
	KALKSTEIN and LIBBY	G.M. counters	Negative	$> 2 \cdot 10^{17}$		
	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 2 \cdot 10^{15}$		
	FIREMAN and SCHWARTZER	Could chamber	Negative	$> 10^{17}$		
	PEARCE and DARBY	Scintillation counters	Negative	$> 4.5 \cdot 10^{16}$		
	MC CARTHY	Scintillation counters	Negative	$> 1.5 \cdot 10^{17}$		
$^{128}_{52}Te \rightarrow ^{128}_{54}Xe$	INGHRAM and REYNOLDS	Chemical separation	Negative	$> 8 \cdot 10^{19}$	$4.4 \cdot 10^{18\pm2}$	$4 \cdot 10^{25\pm2}$
	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 1.3 \cdot 10^{16}$		
$^{130}_{52}Te \rightarrow ^{130}_{54}Xe$	INGHRAM and REYNOLDS	Chemical separation	Positive?	$1.4 \cdot 10^{21}$	$7.4 \cdot 10^{15\pm2}$	$5 \cdot 10^{20\pm2}$
	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 1.3 \cdot 10^{16}$		
	HAYDEN and INGHRAM	Chemical separation	Positive?	$3.3 \cdot 10^{21}$		
$^{150}_{60}Nd \rightarrow ^{150}_{62}Sm$	COWAN *et al.*	Scintillation counters	Negative	$> 6 \cdot 10^{18}$	$2 \cdot 10^{15\pm2}$	$1.5 \cdot 10^{19\pm2}$
$^{238}_{92}U \rightarrow ^{238}_{94}Pu$	LEVINE *et al.*	Chemical separation	Negative	$> 2 \cdot 10^{18}$	$8.5 \cdot 10^{17\pm2}$	$2 \cdot 10^{24\pm2}$

(only for $^{64}_{30}Zn \rightarrow {}^{64}_{28}Ni$).

Transition	Authors	Experimental method	Experimental result	Experimental half-life	Theor. half-life (years)	
					Without neutrinos	With neutrinos
$^{50}_{23}Cr \rightarrow {}^{50}_{22}Ti$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 2.2 \cdot 10^{17}$	∞	∞
$^{54}_{26}Fe \rightarrow {}^{54}_{24}Cr$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 1.3 \cdot 10^{17}$	∞	∞
$^{58}_{28}Ni \rightarrow {}^{58}_{26}Fe$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 3.2 \cdot 10^{17}$	∞	∞
$^{64}_{30}Zn \rightarrow {}^{64}_{28}Ni$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 2 \cdot 10^{17}$	∞	∞
$^{64}_{30}Zn \rightarrow {}^{64}_{28}Ni$ (double electron capture)	BERTHELOT *et al.*	Proportional counters	Negative	$> 8 \cdot 10^{17}$	$1.6 \cdot 10^{26 \pm 2}$	$2 \cdot 10^{28 \pm 2}$
$^{84}_{38}Sr \rightarrow {}^{84}_{36}Kr$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 3.4 \cdot 10^{17}$	∞	∞
$^{92}_{42}Mo \rightarrow {}^{92}_{40}Zr$	FREMLIN and WALTERS	Nuclear emulsion	Positive?	$1.5 \cdot 10^{16}$	∞	∞
	WINTER	Cloud chamber	Negative	$> 4 \cdot 10^{18}$		
$^{106}_{46}Cd \rightarrow {}^{106}_{48}Pd$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 10^{17}$	$2 \cdot 10^{20 \pm 2}$	$1 \cdot 10^{28 \pm 2}$
	WINTER	Cloud chamber	Negative	$> 6 \cdot 10^{16}$		
$^{130}_{56}Ba \rightarrow {}^{130}_{54}Xe$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 1.8 \cdot 10^{15}$	$7 \cdot 10^{20 \pm 2}$	$1 \cdot 10^{29 \pm 2}$
$^{132}_{56}Ba \rightarrow {}^{132}_{54}Xe$	FREMLIN and WALTERS	Nuclear emulsion	Negative	$> 1.8 \cdot 10^{15}$		

cedures. It is clear, on the other hand, that the only conclusive experimental result would be to prove or disprove the existence of a double β-decay in which the sum of the electron energies is equal to the energy release (apart from the recoil energy of the nucleus). It should however be noted that, even if in all β-decays to be found the sum of the electron energies were smaller than the energy release, this would rule out a Majorana theory, but would represent no evidence for the conservation of neutrinos. From a less pessimistic experimental point of view it should be noted that no experimental evidence has been obtained so far for the *existence* of double β-decay; thus the « real » discovery of this type of decay regardless of its nature, would be by itself of considerable interest.

* * *

We wish to than Prof. P. CALDIROLA and G. POLVANI for their interest in this work and Dr. CARLA FIORINI MANGILI for help in collection of the data. One of us (G.F.D.) would like to thank the N.S.F. for a grant and Prof. A. J. F. SIEGERT for his hospitality at Northwestern University.

REFERENCES

[1] W. PAULI: *Nuovo Cimento*, **6**, 204 (1957).
[2] For the latest proposal in this direction see E. T. LEE and C. N. YANG: *Implication of the intermediate boson basis of weak interactions*, Institute for Advanced Studies, preprint; see also T. D. LEE and C. N. YANG: *Possible high energy neutrinos experiments*, preprint and talk delivered at the New York Meeting of the A.P.S., January, 1960.
[3] S. S. KAHANA and D. L. PURSEY: *Nuovo Cimento*, **6**, 1469 (1957).
[4] T. D. LEE and C. N. YANG: *Phys. Rev.*, **105**, 1671 (1957).
[5] L. LANDAU: *Nucl. Phys.*, **3**, 127 (1957).
[6] A. SALAM: *Nuovo Cimento*, **5**, 299 (1957).
[7] C. D. ENZ: *Nuovo Cimento*, **6**, 250 (1957).
[8] G. LÜDERS: *Nuovo Cimento*, **7**, 171 (1958).
[9] M. G. MAYER and V. L. TELEGDI: *Phys. Rev.*, **107**, 1445 (1957).
[10] M. A. PRESTON: *Canad. Journ. Phys.*, **35**, 1017 (1957).
[11] P. E. CAVANAGH, C. F. COLEMAN, B. W. RIDLEY and J. F. TURNER: *Nucl. Phys.*, **5**, 11 (1958).
[12] R. GATTO: *Nucl. Phys.*, **5**, 530 (1958).
[13] R. P. FEYNMAN and M. GELL-MANN: *Phys. Rev.*, **109**, 193 (1958).
[14] E. C. G. SUDARSHAN and R. E. MARSHAK: *Phys. Rev.*, **109**, 1860 (1958); and *Procedings of the Padua-Venice Conference* (Padua, 1958) V-14.
[15] F. GURSEY: *On the structure and parity of weak interaction current*, Institute of Advanced Studies, preprint.

[16] R. Feynman, M. Gell-Mann and M. Lévy: *The axial vector current in β-decay*, preprint.
[17] Y. Nambu: *Phys. Rev. Lett.*, **4**, 380 (1960).
[18] H. Primakoff and S. P. Rosen: *Rep. Prog. Phys.*, **12**, 121 (1959) (also for previous references).
[19] F Yanoukh: *Journ. Exp. Theor. Phys.*, **36**, 335 (1959); *Žurn. Èksp. Teor. Fiz.*, **9**, 231 (1959).
[20] F. Reines, C. L. Cowan Jr., F. B. Harrison, A. D. McGuire and H. W. Kruse: *Phys. Rev.*, **117**, 159 (1960) (also for previous references).
[21] C. P. Enz: *Helv. Phys. Acta*, **31**, 69 (1958).
[22] M. G. Mayer: *Phys. Rev.*, **48**, 512 (1935).
[23] H. Fulbright: *β- and γ-ray Spectroscopy*, ed. Siegbahn (New York, 1955), p. 668.
[24] J. Blatt and V. Weisskopf: *Theoretical Nuclear Physics* (New York, 1952).
[25] E. J. Konopinski and H. Mahmoud: *Phys. Rev.*, **92**, 1045 (1953).
[26] E. L. Fireman: *An experiment on the double β-decay*, Unpublished Doctor's Thesis (Princeton, 1948).
[27] E. L. Fireman: *Phys. Rev.*, **74**, 1238 (1948).
[28] E. L. Fireman: *Phys. Rev.*, **75**, 323 (1949).
[29] M. I. Kalkstein and W. F. Libby: *Phys. Rev.*, **85**, 368 (1952).
[30] A. Berthelot, R. Chaminade, C. Levi and L. Papineau: *Compt. Rend.*, **236**, 1769 (1953).
[31] H. Selig: Unpublished Doctor's Thesis (Carnegie Institute of Technology, 1954).
[32] H. Selig: A.E.C. Report, NYO-3626 (1954).
[33] R. M. Pearce and E. K. Derby: *Phys. Rev.*, **86**, 1049 (1952).
[34] H. W. Fullbright: *Physica*, **18**, 1026 (1952).
[35] J. A. McCarthy: *Phys. Rev.*, **90**, 853 (1953).
[36] J. A. McCarthy: *Phys. Rev.*, **97**, 1234 (1955).
[37] M. Awschalom: *Phys. Rev.*, **101**, 1041 (1956).
[38] J. F. Detoeuf and R. Moch: *Journ. Phys. Rad.*, **16**, 897 (1955).
[39] C. L. Cowan, F. B. Harrison, L. M. Langer and F. Reines: *Nuovo Cimento*, **3**, 649 (1956).
[40] C. L. Cowan and F. Reines: *Phys. Rev.*, **106**, 825 (1957).
[41] E. I. Dobrokhotov, W. R. Lazarenko and S. Yu. Luk'ianov: *Doklady Akad. Nauk. U.S.S.R.*, **110**, 966 (1956); *Soviet Phys. Doklady*, **1**, 600 (1956).
[42] E. I. Dobrokhotov, W. R. Lazarenko and S. Yu. Luk'ianov: *Žurn. Èksp. Teor. Fis.*, **36**, 76 (1959); *Soviet Physics, JETP*, **9**, 54 (1959).
[43] Ya. B. Zel'dovich, S. Yu. Luk'ianov and Ya. A. Smorodinskii: *Usp. Fiz. Nauk.*, **54**, 361 (1954).
[44] K. W. Marlow: *Bull. Am. Phys. Soc.*, **5**, 37 (1960).
[45] J. S. Lawson: *Phys. Rev.*, **81**, 299 (1951).
[46] E. L. Fireman and D. Schwartzer: *Phys. Rev.*, **86**, 451 (1952).
[47] R. G. Winter: *Phys. Rev.*, **85**, 687 (1952).
[48] R. G. Winter: *Phys. Rev.*, **99**, 88 (1955).
[49] J. H. Fremlin and M. C. Walters: *Proc. Phys. Soc. London*, A **65**, 911 (1952).
[50] M. Reinhartz and G. Vanderhaege: *Nuovo Cimento*, **12**, 243 (1954).
[51] J. H. Fremlin and M. C. Walters: *Proc. Phys. Soc. London*, A **63**, 1178 (1950).
[52] M. C. Walters: *Proc. Phys. Soc. London*, A **65**, 959 (1952).
[53] M. G. Inghram and J. H. Reynolds: *Phys. Rev.*, **76**, 1265 (1949).
[54] M. G. Inghram and J. H. Reynolds: *Phys. Rev.*, **78**, 822 (1950).
[55] R. J. Hayden and M. G. Inghram: *Nat. Bur. Stand. Circ.*, **522**, 189 (1953).

[56] C. A. LEVINE, A. GHIORSO and G. T. SEABORG: *Phys. Rev.*, **77**, 296 (1950).
[57] T. L. COLLINS, A. O. NIER and W. H. JOHNSON: *Phys. Rev.*, **86**, 408 (1952).
[58] J. T. MCCARTHY: *Phys. Rev.*, **95**, 447 (1954).
[59] H. E. DUCKWORTH: *Prog. Nucl. Phys.*, **6**, 138 (1957).
[60] A. H. WAPSTRA: *Physica*, **21**, 367 (1955).
[61] J. MATTAUCH and F. EVERLING: *Prog. Nucl. Phys.*, **6**, 267 (1957).
[62] C. F. GIESE and J. L. BENSON: *Phys. Rev.*, **110**, 712 (1958).
[63] D. STROMINGER, J. L. HOLLANDER and G. T. SEABORG: *Rev. Mod. Phys.*, **30**, 585 (1958).
[64] *Nuclear masses and their determination*, Proceedings of the Conference held in the « Max Plank Institut für Chemie », Mainz, July 10-12, 1956, ed. H. HIRTENBURGER (New York, 1957).
[65] W. H. JOHNSON Jr. and A. O. NIER: *Phys. Rev.*, **105**, 1015 (1957).
[66] B. VAN NOOIJEN, J. KONIJN, A. HEYLINGERS, J. F. VAN DER BRUGGE and A. H. WAPSTRA: *Physica*, **23**, 753 (1957).
[67] K. E. EKLUND and R. D. BENT: *Phys. Rev.*, **112**, 966 (1958).
[68] V. B. BELIAEV and B. N. ZAKHAR'EV: *Žurn. Èksp. Teor. Fiz.*, **34**, 505 (1958); *Soviet Physics JETP.*, **7**, 347 (1958).
[69] L. V. GUSTOVA and O. V. CHUBINSKII: *Žurn. Èksp. Teor. Fiz.*, **35**, 1369 (1958), *Soviet Physics JETP.*, **8**, 957 (1959).
[70] P. H. STELSON and F. K. MCCOWAN: WASH (Brookhaven National Laboratory. Associated Universities Inc.), 1018 (1959).
[71] B. S. DZHELEJOV, V. A. ELISEEV, V. P. PRIKHODTSEVA and YU. V. KHOLVOV: *Izv. Akad. Nauk. SSSR, Ser. Fiz.*, **23**, 207 (1959).
[72] N. P. HEYDENBURG, G. F. PIEPER and C. E. ANDERSON: *Phys. Rev.*, **108**, 106 (1957).
[73] V. A. NIEREMBURG: U.C.R.L.-8553 (1958).
[74] R. L. ROBINSON and R. M. LANGER: *Phys. Rev.*, **112**, 481 (1958).
[75] F. BOEHM: *Zeits. f. Phys.*, **152**, 384 (1958).
[76] R. P. SCHUMANN: I.D.O. (Phillips Petroleum Co. Atomic Energy Division, Idaho Falls, Idaho), 16543, 35 (1959).
[77] E. G. FUNK and M. L. WIEDENBECK: *Phys. Rev.*, **112**, 1247 (1958).
[78] H. W. TAYLOR and S. A. SCOTT: *Phys. Rev.*, **114**, 121 (1959).
[79] A. C. G. MITCHELL, J. O. JULIANO, C. B. KREAGER and C. W. KOCHER: *Phys. Rev.*, **113**, 628 (1959).
[80] H. L. GARVIN, T. M. GREEN and E. LIPWORTH: *Phys. Rev. Lett.*, **1**, 292 (1958).
[81] J. E. SHERWOOD and S. J. OVENSHINE: *Phys. Rev.*, **114**, 858 (1959).
[82] W. G. SMITH, P. H. STELSON and F. K. MCGOWAN: *Phys. Rev.*, **114**, 1345 (1959).
[83] R. H. GIRGIS and R. VAN LIESHOUT: *Nucl. Phys.*, **12**, 204 (1959).
[84] N. R. JOHNSON and G. D. O'KELLEY: O.R.N.L., 2782 (1959).
[85] H. RYDE and C. J. HERRLANDER: *Ark. f. Fys.*, **13**, 177 (1958).
[86] M. H. GRACE, C. E. JOHNSON, R. G. SANLOCK and R. T. TAYLOR: *Phil. Mag.*, **3**, 456 (1958).
[87] S. K. BHATTACHERJEE, B. SAHAI and C. V. K. BABA: *Nucl. Phys.*, **12**, 356 (1959).
[88] L. F. KALINKIN, A. S. MELIORANSKII and I. U. ESTULIN: *Žurn. Èksp. Teor. Fiz.*, **36**, 1613 (1959); *Soviet Physics JETP.*, **9**, 1146 (1959).
[89] W. C. PORTER, D. M. VAN PATTER, M. A. ROTHMANN and C. E. MANDEVILLE: *Phys. Rev.*, **112**, 468 (1958).
[90] E. R. BAUMINGER and S. G. COHEN: *Phys. Rev.*, **110**, 953 (1958).
[91] H. DANIEL: *Zeits. f. Phys.*, **150**, 144 (1958).
[92] R. R. SPENCER, G. C. PHYLLIPS and T. E. YUNG: *Phys. Rev.*, **108**, 69 (1957).

CERN

[93] M. Rozkos: *Czechoslov. Journ. Phys.*, **7**, 499, (1957).
[94] W. D. Schmidt: *Zeits. f. Phys.*, **154**, 286 (1959).
[95] R. M. Sinclair: *Phys. Rev.*, **107**, 1306 (1957).
[96] R. P. de Figuerreido, M. Mazzari and W. W. Bruchner: *Phys. Rev.*, **112**, 873 (1958).
[97] D. J. Horen, W. E. Meyerhof, J. J. Kraushaar, D. O. Wells, E. Brun and J. E. Neighbor: *Phys. Rev.*, **113**, 875 (1959).
[98] G. Ythier, W. Schoo, B. L. Schram, H. L. Polak, R. K. Girgis, R. A. Ricci and R. Van Lieshout: *Physica*, **25**, 694 (1959).
[99] R. K. Girgis and R. Van Lieshout: *Physica*, **25**, 688 (1959).
[100] R. B. Duffield and S. H. Vegors: *Phys. Rev.*, **112**, 1958 (1958).
[101] Yu. A. Nemilov, A. N. Pisarevskii and L. D. Soshin: *Izv. Akad. Nauk. S.S.S.R., Sez. Fiz.*, **23**, 255 (1959).
[102] G. D. O'Kelley and N. R. Johnson: O.R.N.L., 2386, 9 (1957).
[103] J. Konijn, H. L. Hagedoorn, H. Van Krugten and J. Slobben: *Physica*, **24**, 931 (1958).
[104] L. Cohen: *Phys. Rev.*, **111**, 587 (1958).
[105] H. I. West jr., L. G. Mann and G. M. Iddings: *Phys. Rev.*, **113**, 881 (1959).
[106] W. B. Ewbank, W. A. Nierember, H. A. Shugart and H. B. Silsbee: *Phys. Rev.*, **110**, 595 (1958).
[107] J. B. Reynolds, R. L. Christensen, D. R. Hamilton, W. M. Hooke and H. H. Stroke: *Phys. Rev.*, **109**, 465 (1958).
[108] W. B. Ewbank, L. L. Marino, W. A. Nieremberg, H. A. Shugart and S. H. Silsbee: *Phys. Rev.*, **115**, 614 (1959).
[109] R. L. Robinson, F. K. McGowan and W. G. Smith: O.R.N.L., 2718 (1959).
[110] R. K. Girgis and R. Van Lieshout: *Physica*, **25**, 1200 (1959).
[111] R. K. Girgis and R. Van Lieshout: *Physica*, **25**, 597 (1959).
[112] D. G. Alkhazov, D. S. Andreev, K. I. Eronkhina and I. Kh. Lemberg: *Žurn. Ėkps. Teor. Fiz.*, **33**, 1347 (1957).; *Soviet Physics JETP.*, **6**, 1036 (1957).
[113] C. L. McGinnis: *Phys. Rev.*, **109**, 888 (1958).
[114] D. S. Hanner and M. L. Perlman: *Phys. Rev.*, **114**, 1133 (1959).
[115] R. K. Girgis and R. Van Lieshout: *Physica*, **25**, 133 (1959).
[116] K. Sugiyama, H. Ikegami, T. Yamazaki and M. Sakai: *Journ. Phys. Soc. Jap.*, **14**, 986 (1959).
[117] H. L. Carvin, T. M. Green and E. Lipworth: *Phys. Rev.*, **111**, 534 (1958).
[118] S. Ofer: *Phys. Rev.*, **113**, 895 (1959).
[119] A. G. W. Cameron: *Chalk River Project Report*, 690 (1957).
[120] F. T. Porter and P. P. Day: *Phys. Rev.*, **114**, 1286 (1959).

891

SUPPLEMENTO AL VOLUME XVII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1960
3° Trimestre

Linear Vector Spaces with Indefinite Metric.

L. K. PANDIT
Tata Institute of Fundamental Research - Bombay

(*Suppl. Nuovo Cimento*: **11**, 157 (1959))

ERRATUM

In the paper with the above title there is a mathematical error on page 171 In the case a real eigenvalue p of a self adjoint operator P occurs as a non-simple root of the minimum polynomial $m(\lambda)$ of P,

$$m(\lambda) = (\lambda - p)^k \cdot q(\lambda), \qquad q(p) \neq 0\,, \qquad k > 1\,,$$

we say that *every* eigenvector χ of P, corresponding to this eigenvalue p, must have the representation

$$\chi = (P - PI)^{k-1}\varphi\,,$$

where φ is a suitable non-zero vector. This statement is wrong. The error lies in concluding that

$$(P - pI)\chi \neq (P - pI)^k\varphi\,, \qquad \text{(for any } \varphi\text{)},$$

if

$$\chi \neq (P - pI)^{k-1}\varphi\,, \qquad \text{(for any } \varphi\text{)}.$$

This is wrong since $(P - pI)$ is a *singular* operator.

Thus the correct statement is that, in the case considered, *at least one* (and not every) eigenvector χ has the representation

$$\chi = (P - pI)^{k-1}\varphi\,,$$

with a suitable φ, and hence

$$\|\chi\| = 0\,.$$

Other eigenvectors may or may not have norm zero.

In the special case of two dimensions our conclusion and its applications (Sect. 2, § 1, 3) later remain unchanged, since in this case when a self-adjoint operator is non-diagonalizable, it has only one eigenvector and so it must be the one with norm zero.

* * *

I wish to thank Mr. N. MUKUNDA for having kindly drawn my attentation to this error.

PROPRIETÀ LETTERARIA RISERVATA

Direttore responsabile: G. POLVANI Tipografia Compositori - Bologna

Questo fascicolo del *Supplemento* è stato licenziato dai torchi il 14-X-1960

893